PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/90, 15/86, 15/31, 15/48, 15/62, 7/01, 9/22, A61K 31/70, 39/21, 48/00

A1

(11) International Publication Number: **WO 96/37626**

(43) International Publication Date: 28 November 1996 (28.11.96)

(21) International Application Number: PCT/US96/06727

(22) International Filing Date: 10 May 1996 (10.05.96)

(30) Priority Data:
08/445,466 22 May 1995 (22.05.95) US

(71) Applicants: CHIRON VIAGENE, INC. [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd Floor, 300 Lakeside Drive, Oakland, CA 94612-3550 (US).

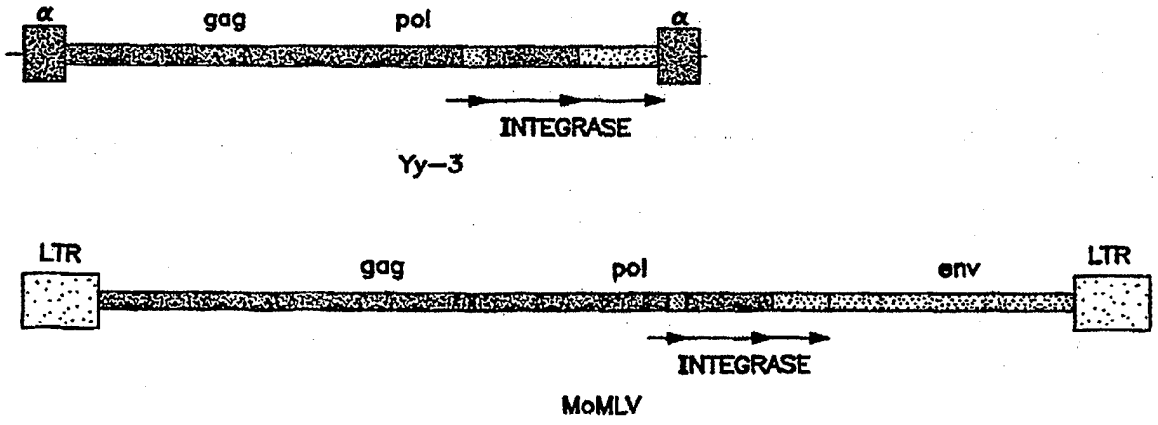
(72) Inventors: RESPESS, James, G.; 4966 Lamont Street, San Diego, CA 92109 (US). SANDMEYER, Suzanne, B.; 226 Jasmine, Corona Del Mar, CA 92625 (US). JOLLY, Douglas, J.; 277 Hillcrest Drive, Encinitas, CA 92024 (US). BILANCHONE, Virginia, W.; 13 Sundown Pass, Irvine, CA 92714 (US). DILDINE, Sandra, L.; 26 Trawler, Laguna Niguel, CA 92677 (US).

(74) Agents: KRUSE, Norman, J. et al.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: POSITION-SPECIFIC INTEGRATION OF VECTOR CONSTRUCTS INTO EUKARYOTIC GENOMES MEDIATED BY A CHIMERIC INTEGRASE PROTEIN



(57) Abstract

Compositions and methods are provided for directing the position-specific integration of a vector construct encoding one or more desired genes into a specific region of a target eukaryotic genome. Such compositions and methods are useful for performing somatic and germ cell gene therapy. Among the benefits provided by position-specific integration are more uniform levels of expression of a desired gene(s) and substantial elimination of the potential for insertional mutagenesis.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

POSITION-SPECIFIC INTEGRATION OF VECTOR CONSTRUCTS INTO EUKARYOTIC GENOMES MEDIATED BY A CHIMERIC INTEGRASE PROTEIN

Technical Field

The present invention relates generally to the field of gene therapy, and more specifically, to compositions and methods useful for introducing vector constructs into specific regions of target eukaryotic cell genomes.

Background Of The Invention

Since the discovery of DNA as the genetic material in the 1940s and continuing through the most recent era of recombinant DNA technology, substantial research has been undertaken in order to realize the possibility that the course of disease may be affected through interaction with the nucleic acids of living organisms. Most recently, a wide variety of methods have been described for altering or affecting genes, including for example, viral vectors derived from retroviruses, adenoviruses, vaccinia viruses, herpes viruses, and adeno-associated viruses (see Jolly, *Cancer Gene Therapy 1*(1):51-64, 1994), as well as physical methods of gene transfer such as lipofection (Felgner et al., *Proc. Natl. Acad. Sci. USA 84*:7413-7417, 1989), direct DNA injection (Acsadi et al., *Nature 352*:815-818, 1991), microprojectile bombardment (Williams et al., *PNAS 88*:2726-2730, 1991), liposomes of several types (*see, e.g.*, Wang et al., *PNAS 84*:7851-7855, 1987) and administration of nucleic acids alone (WO 90/11092).

Of the techniques studied to date, recombinant retroviral gene delivery methods have been most extensively utilized, in part due to: (1) the efficient entry of genetic material (the vector genome) into cells; (2) an active, efficient process of entry into the target cell nucleus; (3) relatively high levels of gene expression; (4) the potential to target particular cellular subtypes through control of the vector-target cell binding and the tissue-specific control of gene expression; (5) a general lack of pre-existing host immunity; and (6) substantial knowledge and clinical experience which has been gained with such vectors.

Briefly, retroviruses are diploid positive-strand RNA viruses that replicate through an integrated DNA intermediate. Typically, retroviruses comprise a protein-containing lipid envelope surrounding a protein-encapsidated core carrying the viral genome. Retroviral infection is initiated by attachment of a viral particle to a specific receptor on the surface of a eukaryotic cell, after which the cell and viral membranes fuse, releasing the genome-containing capsid into the cell's cytoplasm. The retroviral genome is then reverse-transcribed into double stranded linear DNA by a virally encoded reverse transcriptase enzyme that is among the components packaged into each infectious

retroviral particle. Following processing of the linear DNA by the integrase (IN) protein (also encoded by the retrovirus and packaged into the retroviral particle along with reverse transcriptase during virion assembly) to remove two nucleotides from each 3' end of the DNA, the nucleoprotein complex enters the nucleus, where the viral DNA is then integrated pseudo-randomly by staggered cleavage of the host cell genome and joining of the viral DNA to the host cell DNA by strand transfer, creating a "provirus" form of the retroviral genome which is inherited by daughter cells. The processing and joining reactions are mediated by IN and the position of pseudo-random integration appears to be determined mainly by host DNA accessibility, not by nucleotide sequence (Kulkosky, et al. (1994), Pharmac. Ther., vol. 61:185-203).

Wild-type retroviral genomes (and their proviral copies) contain at least three genes (the *gag*, *pol* and *env* genes), which are preceded by a packaging signal (y), and long terminal repeat (LTR) sequences which flank both ends. Briefly, the *gag* gene encodes the internal core structural proteins. The *pol* gene codes for the RNA-dependent DNA polymerase which reverse transcribes the RNA genome and for IN, and the *env* gene encodes the retroviral envelope glycoproteins. The 5' and 3' LTRs contain *cis*-acting elements necessary to promote transcription and polyadenylation of retroviral RNA.

Adjacent to the 5' LTR are sequences necessary for reverse transcription of the genome (the tRNA primer binding site) and for efficient encapsidation of retroviral RNA into particles (the y sequence). Removal of the packaging signal prevents encapsidation (packaging of retroviral RNA into infectious virions) of genomic RNA, although the resulting mutant can still direct synthesis of all proteins encoded in the viral genome.

Recombinant retroviruses and various uses thereof have been described in numerous references including, for example, Mann et al. (*Cell 33:*153, 1983), Cane and Mulligan (*Proc. Nat'l. Acad. Sci. USA 81*:6349, 1984), Miller et al., *Human Gene Therapy* 1:5-14, 1990, U.S. Patent Nos. 4,405,712; 4,861,719; 4,980,289, PCT Application Nos. WO 89/02,468; WO 89/05,349 and WO 90/02,806, and United States patent application serial nos. 08/116,827 (filed September 3, 1993), 08/116,828 (filed September 3, 1993), 08/116,983 (filed September 3, 1993), 08/366,851 (filed December 30, 1994), 08/425, 180 (filed April 20, 1995), and 08/425,762 (filed April 20, 1995). Briefly, one or more foreign genes of interest may be incorporated into the retrovirus in place of a large portion of the normal retroviral RNA. The resultant recombinant retroviral vector constructs are packaged into infectious retroviral particles using one of several systems to supply the requisite retroviral proteins needed to form infectious virions, *i.e.*, the various proteins encoded by the *gag*, *pol*, and *env* genes. Following infection of a cell with such a recombinant retrovirus, the recombinant genome encoding

the foreign gene(s) may then be integrated into the host's cellular DNA as if it were the retrovirus itself. Expression of this foreign gene(s) within the host results in expression of the foreign protein by the host cell.

Despite the utility of recombinant retroviruses as gene delivery vehicles, several shortcomings are known, including their general ability to infect only replicating cells, their inherent genome packaging size limitations, the potential for insertional mutagenesis, *i.e.*, oncogene activation or disruption of a tumor suppressor gene or other essential gene required for cell survival, due to the random nature of retroviral integration, and the potential for contamination of recombinant retroviral preparations with wild-type, replication competent retrovirus (RCR) which may arise as a result of one or more recombination events between vectors used to produce the recombinant virions in various packaging or producer cell systems.

Due to the apparent ability of retroviruses to infect only replicating cells, methods for increasing the efficacy of recombinant retroviruses have been developed. Such methods typically aim at inducing cells to replicate, thereby allowing the retroviruses to infect the cells. Such methods have included, for example, chemical treatment with 10% carbon tetrachloride in mineral oil (Kaleko et al., *Human Gene Therapy 2*:27-32, 1991) and surgery to remove particular tissue, thereby stimulating rapid cell division to increase cell infectability (Rettinger et al., *PNAS 91*:1460-1464, 1994; Moscioni et al., *Surgery 113*:304-311, 1993;). In addition, other methods for transducing cells (*i.e.*, infecting cells with a recombinant retrovirus) resistant to standard transduction techniques, such as stem cells and non-dividing cells, using preparations of high titer recombinant retroviral particles substantially free from contamination with replication competent retrovirus, have recently been developed. *See* U.S.S.N. 08/425,180, *supra*.

Concerns about contamination of recombinant retroviral preparations with RCR have been substantially avoided or eliminated by the development of a variety of packaging cell systems designed to prevent RCR production. Such systems are described in greater detail below.

Despite these and other advances in retroviral gene delivery technology, some scientists have suggested that other potentially more efficient methods of gene transfer, such as direct administration of pure plasmid DNA (Davis et al., *Human Gene Therapy 4*:733-740, 1993) or, alternatively, other viral gene delivery vehicles, for instance those based on DNA viruses or non-integrating RNA viruses, be utilized. However, these alternative systems also have shortcomings, such as providing for only transient expression, thereby requiring repeat administration for treatment of many diseases amenable to a gene therapy approach although, due to pre-existing immunity of many potential patients to the wild-type forms of such alternative viral systems or, in the case of

naked nucleic acids, to the immunogenicity of the gene delivery vehicles themselves, repeated therapy using the same gene delivery vehicle may not be possible.

In view of the current state of the art with respect to gene delivery vehicles, there exists a need to provide gene delivery vehicles capable of providing stable, long term expression of one or more desired genes without many of the shortcomings of the existing gene delivery technologies. Toward this end, it is the object of the present invention to provide compositions and methods enabling the position-specific introduction of a vector construct encoding one or more desired genes into a eukaryotic genome. Accordingly, the present invention has implications in the areas of somatic cell and germ cell gene therapy.

Summary of the Invention

Briefly stated, the present invention provides chimeric integrase proteins capable of directing integration of a vector construct into a defined region of a target eukaryotic genome, as well as methods for making and utilizing chimeric integrases proteins. Within one embodiment of this aspect, the chimeric integrase protein directs integration of a vector construct into a region adjacent to a eukaryotic gene transcribed by RNA polymerase III. Preferably, such integration typically occurs within less than about 1,000 bp of the RNA pol III transcription initiation site, within less than about 100 bp of the transcription initiation site being more preferred, and within less than 10 bp, *i.e.*, 9, 8, 7, 6, 5, 4, 3, 2, and 1 bp, of the transcription initiation site being most preferred.

Within one embodiment of this aspect, the position-specificity of the chimeric integrase protein is mediated by a domain from Ty3 integrase. In another embodiment, the chimeric integrase is derived from a retroviral integrase protein. In a preferred embodiment, the chimeric integrase protein is derived from Moloney murine leukemia virus. A representative example of such a chimeric retroviral integrase is one which comprises, from amino to carboxy terminus, an A domain, a B domain, and a C domain, wherein at least one such domain is derived from Ty3 integrase. Specific examples of such chimeric integrases include those selected from the group consisting of AmBmCt, AmBtCm, AmBtCt, AtBtCm, and AtBmCm, wherein "m" denotes a MoMLV integrase derivation and "t" denotes a Ty3 integrase derivation. In other embodiments, the chimeric integrase protein may be isolated or purified.

Yet another embodiment of this aspect relates to a chimeric integrase protein which directs integration of a vector construct into a defined region of a target eukaryotic genome being incorporated into gene delivery vehicle, wherein the gene delivery vehicle further comprises a vector construct encoding a heterologous gene

product, the heterologous gene product being selected from the group consisting of a polypeptide, an antisense RNA, a sense RNA, and a ribozyme.

Within another aspect, vector constructs which encode a chimeric integrase protein of the invention are provided. Such constructs, also referred to as "chimeric integrase expression vectors," comprise at least one element which controls gene expression, *e.g.*, a prokaryotic or eukaryotic transcriptional promoter, an enhancer or locus defining element(s), or other elements which control gene expression by other means such as alternate splicing, nuclear RNA export, post-transcriptional modification of messenger RNA, or post-translational modification of protein, in functional association, *i.e.*, to effect regulation of gene expression, of a gene encoding a chimeric integrase protein.

A related aspect of the invention concerns host cells into which vector constructs which encode a chimeric integrase are introduced, for instance by transformation, transfection, transduction, or any other technique useful for introduction of nucleic acid into a cell. Such host cells include both prokaryotic and eukaryotic host cells. Such cells can be used for various purposes, for instance, to produce chimeric integrase protein. Chimeric integrase may be so produced by culturing host cells under suitable nutrient conditions (which will vary depending upon the host cell employed, the expression system in use, *etc.*) in a manner allowing expression of the chimeric integrase protein. The resultant chimeric integrase may optionally be isolated or purified, for instance by affinity chromatography using an antibody reactive against an epitope of the chimeric integrase. In one embodiment, a chimeric integrase protein so produced is incorporated into a gene delivery vehicle assembled by an *in vitro* process. In another embodiment, such host cells are packaging cells for the production of recombinant viral particles. In addition to the other components required for assembly of infectious viral particle, packaging cells also produce the chimeric integrase protein which is incorporated into the viral particles.

Within another aspect of the invention, gene delivery vehicles are provided which comprise a chimeric integrase protein to direct integration of a vector construct into a defined region of a target eukaryotic genome and the vector construct. One embodiment concerns recombinant retroviral particles comprising a chimeric retroviral integrase protein to direct integration of a recombinant retroviral vector construct into a defined region of a target eukaryotic genome and the recombinant retroviral vector construct. In a preferred embodiment, such recombinant retroviral particles are transduction competent, and more preferably, are substantially free of contamination with replication competent retrovirus. In another preferred embodiment, the defined region of a target eukaryotic genome is a region adjacent to a eukaryotic gene transcribed by RNA

polymerase III. Eukaryotic genes transcribed by RNA pol III include tRNA genes and 5S RNA genes. In another embodiment, transduction competent recombinant retroviral particles are provided which lead to a reduced rate of insertional mutagenesis caused by integration of the recombinant retroviral vector construct into a eukaryotic genome as compared to integration of the recombinant retroviral vector construct mediated by a transduction competent recombinant retroviral particle carrying wild type retroviral integrase protein. Another embodiment involves transduction competent recombinant retroviral particles which, in transduced eukaryotic cells, lead to decreased variation in expression of a gene of interest carried by the recombinant retroviral vector construct as compared to expression of the gene of interest in eukaryotic cells transduced with a transduction competent recombinant retroviral particle carrying wild type retroviral integrase protein.

Within another aspect of the invention, pharmaceutical compositions comprising a gene delivery vehicle according to the invention and a pharmaceutically acceptable carrier are provided. Another embodiment provides pharmaceutical compositions comprising a gene delivery vehicle, for instance a recombinant virus particle, *e.g.*, a recombinant retroviral particle, a recombinant alphavirus particle (for example, a recombinant Sindbis virus particle), a recombinant adenovirus particle, a recombinant adeno-associated virus particle, recombinant herpes virus particle, and a recombinant pox virus particle, that are lyophilized or dehydrated. In a preferred embodiment, the gene delivery vehicle is transduction competent. A particularly preferred embodiment is a transduction competent recombinant retroviral particle that is lyophilized.

Within yet another aspect of the invention, eukaryotic cell genomes (and the corresponding transduced eukaryotic cells) comprising a vector construct integrated into a defined region are provided. Within one embodiment of this aspect, the defined region into which the vector construct is integrated is a region adjacent to a eukaryotic gene transcribed by RNA polymerase III.

Within another aspect of the invention, methods are provided for introducing a vector construct into a eukaryotic cell genome such that there is a reduced rate of insertional mutagenesis caused by integration of the vector construct into the eukaryotic cell genome as compared to the rate of insertional mutagenesis caused by integration of the vector construct by a wild type integrase protein, the method comprising introducing the vector construct into the eukaryotic cell genome using a chimeric integrase protein of the invention.

Within a related aspect, methods are provided for introducing a vector construct into a defined region of a eukaryotic cell genome such that there is decreased variation in expression of a gene of interest from the vector construct in eukaryotic cells

into which the vector construct is introduced as compared to expression of the gene of interest in eukaryotic cells wherein the vector construct is introduced using a wild type integrase protein, the method comprising introducing the vector construct into the eukaryotic cell genome using a chimeric integrase protein of the invention.

Other aspects of the invention relate to methods of treating a disease selected form the group consisting of a genetic disease, a cancer, an infectious disease, a degenerative disease, an inflammatory disease, a cardiovascular disease, and an autoimmune disease. In one embodiment, the methods comprise *in vivo* administration to a patient of a gene delivery vehicle which directs the integration of a vector construct into a defined region of a target eukaryotic genome. Within another embodiment, the methods comprise administering to a patient cells treated *ex vivo* with a gene delivery vehicle which directs the integration of a vector construct into a defined region of a target eukaryotic genome. Such *ex vivo* methods are preferably performed with transduced autologous cells.

Brief Description of the Drawings

Figure 1 is a graphic representation of the Ty-3 and MoMLV genomes.

Figure 2 is a graphic representation of the "A", "B", and "C" domains of the Ty3 and MoMLV integrase proteins. Amino acid numbers are presented at the boundaries of each domain.

Figure 3 is an amino acid alignment between the primary sequences of Ty3 integrase (above) and MoMLV integrase (below) . Boundaries for the "A", "B", and "C" domains are represented by arrows. The conserved H-H-C-C and D-D-E motifs in each protein are shown.

Figure 4 shows the plasmids BAGÆX, pRgpKan, and pgpChimNeo, the latter generically referencing the expression vector used to make the chimeric integrase proteins described in Example 1.

Figure 5 graphically depicts each of the seven chimeric MoMLV/Ty3 integrase proteins.

Figure 6 (a) and (b) illustrate the plasmid constructions used to generate the chimeric integrase proteins described in Example 1.

Figure 7 provides the nucleotide sequences of the oligonucleotides used to generate five of the seven chimeric integrase proteins described in Example 1. Above each oligonucleotide sequence are the complementary sequences of the MLV and Ty3 integrase genes. "//" indicates where the loop-out occurred in each mutagenesis reaction.

Figure 8 is a illustrates the pseudotyping procedure used to produce infectious retroviral particles containing the chimeric integrase proteins in NC10 cells.

Figure 9 illustrates how the integration library described in Example 1 is made.

Detailed Description of the Invention

Prior to setting forth the invention, it may be helpful to an understanding thereof to set forth definitions of certain terms that will be used hereinafter.

A "vector construct" refers to a recombinantly produced nucleic acid molecule which is capable of directing the expression of one or more genes of interest, sometimes referred to herein as "heterologous sequences." The vector construct must include a promoter capable of directing the expression of (*i.e.*, be "functionally associated with") said gene(s) of interest, and a heterologous sequence, the expression product of which is preferably of therapeutic or prophylactic value. The products of the expression of such genes include proteins, polypeptides, antisense RNA, sense RNA, and ribozymes. Optionally, the vector construct may include transcription termination, splice recognition, polyadenylation addition sites, and one or more genes coding for a selectable marker. In a multivalent vector construct, i.e., a vector construct coding for more than one gene of interest, there need not, although there may be, a promoter to direct the expression of each heterologous sequence. In a preferred embodiment wherein the multivalent vector construct codes for the expression of two or more genes of interest, expression of the gene(s) downstream of the first is mediated by an internal ribosome entry site ("IRES") sequence.

"Viral vector", "recombinant viral vector", "viral vector construct", and "recombinant viral vector construct" refers to a nucleic acid construct which carries, and within certain embodiments, is capable of directing the expression of a vector construct according to the invention. The viral vector must include at least one transcriptional promoter/enhancer or locus defining element(s), or other elements which control gene expression by other means such as alternate splicing, nuclear RNA export, post-translational modification of messenger, or post-transcriptional modification of protein. In addition, any other nucleic acid component essential for inclusion within a viral particle of the particular type intended is included. Optionally, the recombinant retroviral vector may also include a signal which directs polyadenylation, selectable markers such as resistance to neomycin, hygromycin, phleomycin, histidinol, or proteins such as DHFR (confers methotrexate resistance) and HSVTK (confers gancilovir sensitivity), as well as one or more restriction sites and a translation termination sequence.

"Retroviral vector construct", "retroviral vector", "recombinant retroviral vector", and "recombinant retroviral vector construct" refers to a nucleic acid construct which carries, and within certain embodiments, is capable of directing the expression of a

vector construct according to the invention. The retroviral vector must include at least one transcriptional promoter/enhancer or locus defining element(s), or other elements which control gene expression by other means such as alternate splicing, nuclear RNA export, post-translational modification of messenger, or post-transcriptional modification of protein. Such vector constructs must also include sequence encoding a packaging signal, long terminal repeats (LTRs) or portion thereof, and positive and negative strand primer binding sites appropriate to the retrovirus used (if these are not already present in the retroviral vector). Optionally, the recombinant retroviral vector may also include a signal which directs polyadenylation, selectable markers such as resistance to neomycin, hygromycin, phleomycin, histidinol, or proteins such as DHFR (confers methotrexate resistance) and HSVTK (confers gancilovir sensitivity), as well as one or more restriction sites and a translation termination sequence. By way of example, such vectors typically include a 5' LTR, a sequence coding for a tRNA binding site, a sequence coding for a packaging signal, an origin of second strand DNA synthesis, and a 3' LTR or a portion thereof.

"Recombinant DNA molecule" is used from time to time in place of "vector construct" or "retroviral vector construct."

A "gene delivery vehicle" refers to a composition capable of delivering a vector construct to an eukaryotic cell which further comprises a chimeric IN protein according to the invention to promote position-specific integration of the vector construct into the genome of the eukaryotic cell. Representative examples of gene delivery vehicles include recombinant viral vectors (e.g. alphaviruses such as Sindbis), physical systems (e.g. ELVS), other viral systems (e.g. adenovirus, adeno-associated virus, and poxvirus), a nucleic acid vector (such as a plasmid), a naked nucleic acid molecule such as genes, a nucleic acid molecule complexed to a polycationic molecule capable of neutralizing the negative charge on the nucleic acid molecule and condensing the nucleic acid molecule into a compact molecule (*see* WO 93/03709), a nucleic acid associated with a liposome (Wang, *et al., PNAS 84*:7851, 1987), a bacterium, and certain eukaryotic cells such as a producer cell, that are capable of delivering a nucleic acid molecule having one or more desirable properties to host cells in an organism. As discussed below, the desirable properties include the ability to express a desired substance, such as a protein, enzyme, or antibody, and/or the ability to provide a biological activity, which is where the nucleic acid molecule carried by the GDV is itself the active agent without requiring the expression of a desired substance. One example of such biological activity is gene therapy where the delivered nucleic acid molecule incorporates into a specified gene so as to inactivate the gene and "turn off" the product the gene was making. Another example is where the nucleic acid sequence is a antisense molecule that binds to mRNA and inhibits translation.

If the nucleic acid sequence encodes a rybozyme the ribozyme binds and cleaves mRNA thereby inhibiting translation.

A "defined region of a target eukaryotic genome" is a region into which a vector construct of the invention is integrated due to the position-specificity of the integrase mediating the integration event. A representative example is a region adjacent to a eukaryotic gene transcribed by RNA polymerase III.

As noted above, the present invention provides compositions and methods for directing the position-specific integration of a vector construct into a eukaryotic cell genome through the use of a chimeric integrase protein. Such compositions are suitable for administration to various eukaryotic organisms, particularly warm-blooded animals, including mammals, and especially humans. By integrating a vector construct into a specific position within a eukaryotic genome, potential obstacles such as cell-to-cell variability in expression levels of the gene(s) of interest and insertional mutagenesis, among others, can be substantially avoided or eliminated. As a result, such compositions will be useful for efficacious gene therapy for a variety of diseases by a variety of routes.

A. PREPARATION OF CHIMERIC INTEGRASE PROTEINS

The enzyme responsible for catalyzing the integration event is the retroviral IN which mediates removal of a 3' dinucleotide from the ends of the extrachromosomal retroviral vector DNA, cleavage of the target site generating a characteristic (4 to 6) bp 5' overhang, and ligation of the 3' ends of the retroviral vector DNA to the 5' ends of the host chromosomal DNA (Fujiwara, et al. (1988), Cell, vol. 54: 497-504; Brown, et al. (1989), Proc. Natl. Acad. Sci. USA, vol. 86: 2525-2529). The MoMLV intracellular viral core particle containing replicated DNA was shown to be competent for integration (Brown, et al. (1987), Cell, vol. 49:347-356; Fujiwara, et al. (1989), Proc. Natl. Acad. Sci. USA, vol. 86:3065-3069). More recently, purified, recombinant IN together with linear molecules representing replicated viral DNA or oligonucleotide duplexes representing the ends of the replicated virus in the presence of buffer and divalent cations was shown to be sufficient to catalyze nicking and ligation, mimicking an integration reaction (Katz, et al. (1990), Cell, vol. 63:87-95; Craigie, et al. (1990), Cell, vol. 62:829-837; Bushman, et al. (1990), Science, vol. 249:1555-1558.). In addition, IN can perform a reversal of the integration reaction on an oligonucleotide substrate DNA (Chow, et al. (1992), Science, vol. 255:723-726).

Ty3 is a yeast retrotransposon which, except for the absence of the *env* gene, is organizationally and functionally similar to animal retroviruses (Hansen, et al.

(1988), Mol. Cell. Biol., vol. 8:5245-5256; Hansen, et al. (1990), J. Virol., vol. 64:2599-2607; Hansen, et al. (1992), J. Virol., vol. 66:1414-1424.). *See* FIG. 1. Ty3 IN is required for nicking of the 3' ends of the replicated Ty3 DNA by two bp and for integration of Ty3. Ty3 is composed of 340 bp LTRs flanking an internal domain of 4.7 kbp. Transcription of the 5.2 kbp genomic RNA begins and ends in the 5' and 3' LTRs, respectively. The internal domain contains two ORFs, *GAG3* and *POL3*, corresponding to the retroviral *gag* and *pol* genes. The *GAG3* gene encodes the precursor polyprotein $Pr38^{GAG3}$, which is processed into a 26 kDa capsid (CA) species and a 9 kDa nucleocapsid (NC) species. These species have conserved motifs found in their retroviral counterparts and are functionally equivalent to those proteins. The $GAG3\text{-}POL3^{173}$ fusion polypeptide is processed to the *GAG3* proteins described above, a 16 kDa aspartyl protease (PR), a p115 *POL3* species composed of reverse transcriptase (RT) which includes IN domains, a 55 kDa RT, and 61 and 58 kDa IN species. These proteins together with Ty3 RNA form virus-like particles (VLPs) about 50 nm in diameter and 156S in size. In addition to having RNA and Ty3 proteins, the particle fraction displays RT activity and contains the full-length, replicated Ty3 DNA. The primer for Ty3 replication is initiator $tRNA^{Met}$ which is complementary to the Ty3 primer binding site which begins 2 bp downstream of the U5 internal domain junction. As is the case for retroviruses, this DNA ends in the short, conserved, inverted repeats and has two bp terminal extensions compared to the integrated form. Thus it is structurally analogous to the retrovirus replicated extrachromosomal intermediate.

In one embodiment of the invention, the position specificity of the yeast Ty3 element is conferred to the integrase (IN) of a Moloney murine leukemia virus-based retroviral vector. Integrase proteins are known to comprise at lease three domains, an amino terminal domain, a core domain ("B" in FIG. 2) with strand transfer and metal chelation activity, while the carboxyl terminal domain ("C" in FIG. 2) is involved in DNA binding. One or more of these discrete domains of an integrase known to confer position-specific integration, *e.g.*, the Ty3 IN, can be substituted for the corresponding domain(s) of the integrase to be incorporated into a gene delivery vehicle according to the invention. Vector constructs of the invention which code for a chimeric integrase carry substitutions of the coding region for the amino terminal region, the core region, and the carboxyl terminal region of Ty3 IN for the analogous coding regions for the nonspecific integrase activity of MoMLV.

The full-length protein sequence of Ty3 IN and MoMLV IN have been aligned (FIG. 3). Mutations in the D-D-E region conserved in Ty3 and in retrovirus IN and lying within region B affect 3' nicking and strand transfer (Kirchner, J. and Sandmeyer, S. (1992). Proteolytic processing of Ty3 proteins is required for

transposition. J. Virol. 67:1, 19-28.Kulkosky, et al. (1992), Mol. Cell. Biol., vol. 12:2331-2338; Engelman, et al. (1992), J. Virol., vol. 66:6361-6369). Although the amino-terminal and carboxyl-terminal regions of IN are required for full 3' processing and strand-transfer activities, these domains also appear to be functionally distinct (Bushman, et al. (1993), Proc. Natl. Acad. Sci. USA, vol. 90:3428-3432; Geiduschek, et al. (1988), Annu. Rev. Biochem., vol. 57:873-914). A subdomain of this core region from HIV expressed as a recombinant protein (Kassavetis, et al. (1990), Cell, vol. 60:235-245) has been shown to be sufficient for mediating a reversal of integration, referred to as a "disintegration" reaction. The conserved acidic residues are likely to chelate the divalent cation and this region contains a domain which encompasses a complete active site.

Seven representative chimeric constructs are described in Example 1, below. Because Ty3 inserts within a few nucleotides of the transcription initiation site and RNA pol III gene sequences are highly conserved and which are not regulated in a tissue-specific manner, predictable and less variable expression levels can be achieved. Northern analysis and PCR are among the techniques that can be employed to assess levels of expression of the gene(s) of interest.

Despite the similarity between the Ty3 IN and its retroviral counterpart, Ty3 integrates with position specificity which is not observed for any retrovirus. Ty3 integrates specifically in the region of transcription initiation of genes transcribed by RNA polymerase III, *e.g.*, 5S, U6, and tRNA genes. The tRNA class of polymerase III transcribed genes is distinguished by the *boxA* and *boxB* internal promoter elements (Chalker, et al. (1990), Genetics, vol. 126:837-850). These regions of the tRNA gene-coding sequence direct binding of transcription factor TFIIIC which subsequently directs binding of TFIIIB to the 5' flanking region of the tRNA gene. TFIIIB bound to the DNA template is sufficient to direct RNA pol III to initiate multiple rounds of transcription (Chalker, et al. (1992), Genes Dev., vol 6:117-128). No consensus DNA binding sequence for TFIIIB has been observed and the region upstream of the gene does not contain conserved promoter elements.

The *in vivo* specificity of Ty3 has been tested directly using a plasmid target assay (Natsoulis, et al. (1989), Genetics, vol. 123:269-279) to showed that tRNA, 5S, and U6 genes are targets of Ty3 integration. Integration events occurred so that the gene-proximal member of the staggered nicks at the integration site were within one or two bp of the position of transcription initiation and immediately downstream of the position of TFIIIB. Ty3 integration into a target tRNA gene requires functional tRNA gene promoter elements (Natsoulis, et al. (1989), Genetics, vol. 123:269-279) and it appears that transcription factors are bound upstream of the gene at the time of integration (Kinsey, et al. (1991), Nucleic Acids Res., vol. 19:1317-1324).

Because RNA pol III promoter elements are internal, integration of recombinant DNA molecules according to the invention by a chimeric integrase having the position-specificity of the Ty3 IN just upstream of the pol III initiation site will not disrupt tRNA gene expression. Analysis of transcription of one tRNA gene alone and with a Ty3 in both orientations under conditions where the level of expression of the Ty3 element or its LTR sequence were varied over 50 fold showed that the Ty3 had only a modest, positive effect on the level of tRNA gene expression (Engelman, et al. (1993), EMBO J., vol. 12:3269-3275). The tRNA genes of eukaryotic cells in which they have been examined are highly redundant, thereby minimizing insertional effects.

Recently an *in vitro* assay has been developed with which to study the components of the Ty3 integration reaction. This reaction is based on the *in vitro* retrovirus core integration assay (Brown, et al. (1987), Cell, vol. 49:347-356). The Ty3 VLP contain, in addition to other components, IN and full-length Ty3 DNA, and can donate the Ty3 DNA to a plasmid target. VLPs are isolated from cells overexpressing Ty3 and mixed with plasmid to which RNA pol III transcription extract or purified factors (IIIC and IIIB) and polymerase III is added. The plasmid contains a modified *SUP2* tRNA gene with either a wild-type or mutant (G56) *boxB* promoter element. The components of the integration reaction are mixed together on ice in buffer with 5-20 mM $MgCl_2$ and then incubated at 30°C for 30 min. The DNA is then extracted and quantitated by a fluorometric assay. Integration of Ty3 DNA into the target plasmid containing the tRNA gene can be monitored by a PCR assay. Integration of the Ty3 DNA into the initiation region of the tRNA gene target results in a product which can be amplified into a diagnostic fragment.

Because of differences in codon usage between eukaryotic retroviruses and retrotransposons, those skilled in the art can modify the respective vector construct using degenerate codons preferred by the particular integrase into which position-specifying domains are to be inserted. Information regarding codon usage in the host(s) susceptible to infection by a retrovirus (or in the cell line to be used for virion or integrase expression) if unavailable in the literature, can be determined by examining codon usage in genes coding for highly expressed proteins. One or more preferred codons can then be introduced into the chimeric integrase gene by standard techniques, such as site directed mutagenesis or solid state nucleic acid synthesis.

A. PREPARATION OF RECOMBINANT RETROVIRAL VECTORS, PACKAGING CELLS, PRODUCER CELLS AND RECOMBINANT RETROVIRUSES

As noted above, one embodiment of the present invention provides recombinant retroviruses which are constructed to deliver one or more selected nucleic

acid molecules, or "genes," of interest in a position-specific manner to the genome of a eukaryotic cell. Position-specific integration is mediated by the chimeric IN protein incorporated into the recombinant retroviral particle. Briefly, numerous retroviral gene delivery vehicles may be utilized within the context of the present invention, including for example those described in EP 0,415,731; WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; U.S. Patent No. 5,219,740; WO 9311230; WO 9310218; Vile and Hart, *Cancer Res. 53*:3860-3864, 1993; Vile and Hart, *Cancer Res. 53*:962-967, 1993; Ram et al., *Cancer Res. 53*:83-88, 1993; Takamiya et al., *J. Neurosci. Res. 33*:493-503, 1992; Baba et al., *J. Neurosurg. 79*:729-735, 1993 (U.S. Patent No. 4,777,127, GB 2,200,651, EP 0,345,242 and WO 91/02805). Particularly preferred recombinant retroviruses include those described in WO 91/02805.

Retroviral gene delivery vehicles of the present invention may be readily constructed from a wide variety of retroviruses, including for example, B, C, and D type retroviruses as well as spumaviruses and lentiviruses (see RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985). Preferred retroviruses for the preparation or construction of retroviral gene delivery vehicles of the present invention include retroviruses selected from the group consisting of Avian Leukosis Virus, Bovine Leukemia Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe, *J. Virol. 19*:19-25, 1976), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998), and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be readily obtained from depositories or collections such as the American Type Culture Collection ("ATCC"; Rockville, Maryland), or isolated from known sources using commonly available techniques.

Any of the above retroviruses may be readily utilized in order to assemble or construct retroviral gene delivery vehicles given the disclosure provided herein, and standard recombinant techniques (*e.g.*, Sambrook et al, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, 1989; Kunkle, *PNAS 82*:488, 1985). In addition, within certain embodiments of the invention, portions of the retroviral gene delivery vehicles may be derived from different retroviruses. For example, within one embodiment of the invention, retroviral vector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

Within preferred embodiments of the present invention, recombinant retroviruses useful in the practice the invention may be made by introducing a vector construct as discussed above, into a cell (termed a "packaging cell") which contains those elements necessary for production of infectious recombinant retrovirus which mediate position-specific integration of the recombinant retroviral genome, but which are lacking in the vector construct from which the recombinant retroviral genome is transcribed. A wide variety of retroviral vector constructs may be utilized within the present invention in order to prepare recombinant retroviruses. For example, within one aspect of the present invention retroviral vector constructs are provided comprising a 5' LTR, a tRNA binding site, a packaging signal, one or more heterologous sequences, an origin of second strand DNA synthesis and a 3' LTR, wherein the vector construct lacks gag/pol or env coding sequences. Briefly, Long Terminal Repeats ("LTRs") are subdivided into three elements, designated U5, R and U3. These elements contain a variety of signals which are responsible for the biological activity of a retrovirus, including for example, promoter and enhancer elements which are located within U3. LTRs may be readily identified in the provirus due to their precise duplication at either end of the genome. As utilized herein, a 5' LTR should be understood to include a 5' promoter element and sufficient LTR sequence to allow reverse transcription and integration of the DNA form of the vector. The 3' LTR should be understood to include a polyadenylation signal, and sufficient LTR sequence to allow reverse transcription and integration of the DNA form of the vector.

The tRNA binding site and origin of second strand DNA synthesis are also important for a retrovirus to be biologically active, and may be readily identified by one of skill in the art. For example, retroviral tRNA binds to a tRNA binding site by Watson-Crick base pairing, and is carried with the retrovirus genome into a viral particle. The tRNA is then utilized as a primer for DNA synthesis by reverse transcriptase. The tRNA binding site may be readily identified based upon its location just downstream from the 5' LTR. Similarly, the origin of second strand DNA synthesis is, as its name implies, important for the second strand DNA synthesis of a retrovirus. This region, which is also referred to as the poly-purine tract, is located just upstream of the 3' LTR.

As position-specific integration is an essential aspect of the invention, it is critical that the retroviral vector construct include sequences which, when positioned in the reverse transcribed double stranded form of the retroviral genome to be integrated, are processed by intact chimeric IN protein. Following reverse transcription of the retroviral RNA genome, LTRs are present at each end of the linear, double stranded DNA molecule. Several nucleotides, typically two, are removed from the 3'-OH ends of the retroviral DNA by the chimeric IN. In a particularly preferred embodiment, the substrate DNA has a cytosine, adenosine (5'-CA-3') dinucleotide two base pairs from the 3'

terminus of each of the two complementary DNA strands. *See* Kulkosky, *et al.*, *supra*. However, resultant retroviral DNAs which contain a substitution of either or both nucleotides or where the dinucleotide is more recessed from one or both ends of the DNA can also be employed, although processing and joining (to the cleaved eukaryotic genome) of the retroviral DNA will likely be less efficient.

In addition to a 5' and 3' LTR, tRNA binding site, and origin of second strand DNA synthesis, certain preferred retroviral vector constructs which are provided herein also comprise a packaging signal, as well as one or more nucleic acid molecules (*e.g.*, heterologous sequences), each of which is discussed in more detail below.

Within a preferred embodiment of the invention, retroviral vector constructs are provided which lack both *gag/pol* and *env* coding sequences. As utilized herein, the phrase "lacks *gag/pol* or *env* coding sequences" should be understood to mean that the retroviral vector does not contain at least 20, preferably at least 15, more preferably at least 10, and most preferably at least 8 consecutive nucleotides which are found in *gag/pol* or *env* genes, and in particular, within *gag/pol* or *env* expression cassettes that are used to construct packaging cell lines for the retroviral vector construct.

Packaging cell lines suitable for use with the above-described retroviral vector constructs may be readily prepared (*see* U.S. Serial No. 08/240,030, filed May 9, 1994; see also WO 92/05266), and utilized to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles which mediate position-specific integration. Within particularly preferred embodiments of the present invention packaging cell lines are derived from cell lines obtained from the same species as will be treated with the resulting recombinant retroviral particles. For instance, it has been discovered that recombinant retroviral particles made from human (*e.g.*, HT1080 cells) cell lines are capable of surviving inactivation in human serum. *See* U.S.S.N. 08/367,071, filed December 30, 1994.

In a preferred embodiment of the invention, packaging cell lines that produce recombinant retroviral particles at titers greater than 10^6 or 10^7 cfu/ml (in crude supernatant) may readily be obtained. In addition, it should be noted that such titers are generally obtained from titer assays on HT1080 cells, which produce a three-fold lower titer than titers obtained on murine 3T3 cells.

B. Recombinant Retroviruses Which Carry and/or Express a Desired Nucleic Acid Molecule

A wide variety of nucleic acid molecules may be carried and/or expressed by the recombinant vector constructs of the present invention. Generally, the nucleic acid molecules which are described herein do not occur naturally in the gene delivery vehicle

that carries it, and provides some desirable benefit, typically an ability to fight or prevent a disease, or other pathogenic agent or condition. As used herein, "pathogenic agent" refers to a cell that is responsible for a disease state. Representative examples of pathogenic agents include tumor cells, autoreactive immune cells, hormone secreting cells, cells which lack a function that they would normally have, cells that have inappropriate gene expression which does not normally occur in that cell type, and cells infected with bacteria, viruses, or other intracellular parasites.

Examples of nucleic acid molecules which may be carried and/or expressed by the gene delivery vehicles of the present invention include genes and other nucleic acid molecules which encode a substance, *e.g.*, a polypeptide, antisense RNA, sense RNA, or a ribozyme, as well as biologically active nucleic acid molecules such as inactivating sequences that incorporate into a specified intracellular nucleic acid molecule and inactivate that molecule. A nucleic acid molecule is considered to be biologically active when the molecule itself provides the desired benefit without requiring the expression of a substance. For example, the biologically active nucleic acid molecule may be an inactivating sequence that incorporates into a specified intracellular nucleic acid molecule and inactivates that molecule, or the molecule may be a tRNA, rRNA or mRNA that has a configuration that provides a binding capability.

Substances which may be encoded by the nucleic acid molecules described herein include proteins (*e.g.*, antibodies, including single chain molecules), immunostimulatory molecules (such as antigens, immunosuppressive molecules, blocking agents, palliatives (such as toxins, antisense ribonucleic acids, ribozymes, enzymes, and other material capable of inhibiting a function of a pathogenic agent) cytokines, various polypeptides or peptide hormones, their agonists or antagonists, where these hormones can be derived from tissues such as the pituitary, hypothalamus, kidney, endothelial cells, liver, pancreas, bone, hematopoietic marrow, and adrenal. Such polypeptides can be used for induction of growth, regression of tissue, suppression of immune responses, apoptosis, gene expression, blocking receptor-ligand interaction, immune responses and can be treatment for certain anemias, diabetes, infections, high blood pressure, abnormal blood chemistry or chemistries (*e.g.*, elevated blood cholesterol, deficiency of blood clotting factors, elevated LDL with lowered HDL), levels of Alzheimer associated amaloid protein, bone erosion/calcium deposition, and controlling levels of various metabolites such as steroid hormones, purines, and pyrimidines.

For palliatives, when "capable of inhibiting a function" is utilized within the context of the present invention, it should be understood that the palliative either directly inhibits the function or indirectly does so, for example, by converting an agent present in the cells from one which would not normally inhibit a function of the pathogenic agent to

one which does. Examples of such functions for viral diseases include adsorption, replication, gene expression, assembly, and exit of the virus from infected cells. Examples of such functions for cancerous diseases include cell replication, susceptibility to external signals (*e.g.*, contact inhibition), and lack of production of anti-oncogene proteins. Examples of such functions for cardiovascular disease include inappropriate growth or accumulation of material in blood vessels, high blood pressure, undesirable blood levels of factors such as cholesterol or low density lipoprotein that predispose to disease, localized hypoxia, and inappropriately high and tissue-damaging levels of free radicals. Examples of such functions for neurological conditions include pain, lack of dopamine production, inability to replace damaged cells, deficiencies in motor control of physical activity, inappropriately low levels of various peptide hormones derived from neurological tissue such as the pituitary or hypothalamus, accumulation of Alzheimer's Disease associated amyeloid plaque protein, and inability to regenerate damaged nerve junctions. Examples of such functions for autoimmune or inflammatory disease include inappropriate production of cytokines and lymphokines, inappropriate production and existence of autoimmune antibodies and cellular immune responses, inappropriate disruption of tissues by proteases and collagenases, lack of production of factors normally supplied by destroyed cells, and excessive or aberrant regrowth of tissues under autoimmune attack.

Within one aspect of the present invention, methods are provided for administration of a gene delivery vehicle, the nucleic acid of which encodes and directs the expression of at least one palliative. Within various embodiments, the palliative may be a DNA molecule, a RNA molecule, some combination of the two, or a protein.

Representative examples of palliatives that act directly to inhibit the growth of cells include toxins such as ricin (Lamb et al., *Eur. J. Biochem. 148*:265-270, 1985), abrin (Wood et al., *Eur. J. Biochem. 198*:723-732, 1991; Evensen et al., *J. of Biol. Chem. 266*:6848-6852, 1991; Collins et al., *J. of Biol. Chem. 265*:8665-8669, 1990; Chen et al., *Fed. of Eur. Biochem Soc. 309*:115-118, 1992), diphtheria toxin (Tweten et al., *J. Biol. Chem. 260*:10392-10394, 1985), cholera toxin (Mekalanos et al., *Nature 306*:551-557, 1983; Sanchez & Holmgren, *PNAS 86*:481-485, 1989), gelonin (Stirpe et al., *J. Biol. Chem. 255*:6947-6953, 1980), pokeweed (Irvin, *Pharmac. Ther. 21*:371-387, 1983), antiviral protein (Barbieri et al., *Biochem. J. 203*:55-59, 1982; Irvin et al., *Arch. Biochem. & Biophys. 200*:418-425, 1980; Irvin, *Arch. Biochem. & Biophys. 169*:522-528, 1975), tritin, Shigella toxin (Calderwood et al., *PNAS 84*:4364-4368, 1987; Jackson et al., *Microb. Path. 2*:147-153, 1987), and Pseudomonas exotoxin A (Carroll and Collier, *J. Biol. Chem. 262*:8707-8711, 1987). A detailed description of recombinant retroviruses which express Russel's Viper Venom is provided in U.S. Serial No. 08/368,574, filed December 30, 1994.

Within other embodiments of the invention, the gene delivery vehicle carries a gene specifying a product which is not in itself toxic, but when processed or modified by a protein, such as a protease specific to a viral or other pathogen, is converted into a toxic form. For example, the recombinant retrovirus could carry a gene encoding a proprotein chain, which becomes toxic upon processing by the HIV protease. More specifically, a synthetic inactive proprotein form of the toxic ricin or diphtheria A chains could be cleaved to the active form by arranging for the HIV virally encoded protease to recognize and cleave off an appropriate "pro" element.

Within yet another embodiment of the invention, the vector construct carried by the gene delivery vehicle directs the expression of a substance capable of activating an otherwise inactive precursor into an active inhibitor of a pathogenic agent, or a conditional toxic palliative, which are palliatives that are toxic for the cell expressing the pathogenic condition. As should be evident given the disclosure provided herein, a wide variety of inactive precursors may be converted into active inhibitors of a pathogenic agent. For example, antiviral nucleoside analogues such as AZT or ddI are metabolized by cellular mechanisms to the nucleotide triphosphate form in order to specifically inhibit retroviral reverse transcriptase, and thus viral replication (Furmam et al., *Proc. Natl. Acad. Sci. USA 83*:8333-8337, 1986). Recombinant retroviruses which direct the expression of a gene product (*e.g.*, a protein) such as Herpes Simplex Virus Thymidine Kinase (HSVTK) or Varicella Zoster Virus Thymidine Kinase (VZVTK) which assists in metabolizing antiviral nucleoside analogues to their active form are therefore useful in activating nucleoside analogue precursors (*e.g.*, AZT or ddC) into their active form. AZT or ddI therapy will thereby be more effective, allowing lower doses, less generalized toxicity, and higher potency against productive infection. Additional nucleoside analogues whose nucleotide triphosphate forms show selectivity for retroviral reverse transcriptase but, as a result of the substrate specificity of cellular nucleoside and nucleotide kinases are not phosphorylated, will be made more efficacious.

Within one embodiment of the invention, the HSVTK gene may be expressed under the control of a constitutive macrophage or T-cell-specific promoter, and introduced into macrophage or T-cells. Constitutive expression of HSVTK results in more effective metabolism of nucleotide analogues such as AZT or ddI to their biologically active nucleotide triphosphate form, and thereby provides greater efficacy, delivery of lower doses, less generalized toxicity, and higher potency against productive infection. Additional nucleoside analogues whose nucleotide triphosphate forms show selectivity for retroviral reverse transcriptase but, as a result of the substrate specificity of cellular nucleoside and nucleotide kinases are not phosphorylated, may also be utilized within the context of the present invention.

Within a related aspect of the present invention, vector constructs are provided which direct the expression of a substance that activates another compound with little or no cytotoxicity into a toxic product in the presence of a pathogenic agent, thereby effecting localized therapy to the pathogenic agent. In this case, expression of the gene product from the recombinant retrovirus is limited to situations wherein an entity associated with the pathogenic agent, such as an intracellular signal identifying the pathogenic state, is present, thereby avoiding destruction of nonpathogenic cells. This cell-type specificity may also be conferred at the level of infection, by targeting the recombinant gene DNA vehicle carrying the vector construct to cells having or being susceptible to the pathogenic condition.

In a related embodiment, vector constructs are provided which direct the expression of a gene product(s) that activates a compound with little or no cytotoxicity into a toxic product. Briefly, a wide variety of gene products which either directly or indirectly activate a compound with little or no cytotoxicity into a toxic product may be utilized within the context of the present invention. Representative examples of such gene products include HSVTK and VZVTK which selectively monophosphorylate certain purine arabinosides and substituted pyrimidine compounds, converting them to cytotoxic or cytostatic metabolites. More specifically, exposure of the drugs gangicylovir, acyclovir, or any of their analogues (*e.g.*, FIAC, DHPG) to HSVTK, phosphorylates the drug into its corresponding active nucleotide triphosphate form. For example, the vector construct may direct the expression of a HSVTK gene under the transcriptional control of an HIV promoter known to be transcriptionally silent except when activated by HIV tat protein. Briefly, expression of the *tat* gene product in human cells infected with HIV and carrying the vector construct would lead to production of HSVTK. The cells (either *in vitro* or *in vivo*) are then exposed to a drug such as ganciclovir, acyclovir or its analogues (FIAC, DHPG). As noted above, these drugs are known to be phosphorylated by HSVTK (but not by cellular thymidine kinase) to their corresponding active nucleotide triphosphate forms. Acyclovir triphosphates inhibit cellular polymerases in general, leading to the specific destruction of cells expressing HSVTK in transgenic mice (see Borrelli et al., *Proc. Natl. Acad. Sci. USA 85*:7572, 1988). Those cells containing the vector construct and expressing HIV tat protein are selectively killed in the presence of an appropriate dose of these drugs.

Within another embodiment of the invention, expression f a conditionally lethal HSVTK gene may be made even more HIV-specific by including *cis*-acting elements in the mRNA transcribed from the vector construct ("CRS/CAR") which require an additional HIV gene product, *e.g.*, *rev*, for optimal activity (Rosen et al., *Proc. Natl. Acad. Sci. USA 85*:2071, 1988). More generally, *cis* elements present in mRNAs have

been shown in some cases to regulate mRNA stability or translatability. Sequences of this type (*i.e.*, for post-transcriptional regulation of gene expression) may be used for event- or tissue-specific regulation of vector gene expression. In addition, multimerization of these sequences (*i.e.*, rev-responsive "CRS/CAR" or tat-responsive "TAR" elements for HIV) may be utilized in order to generate even greater specificity.

In a manner similar to the preceding embodiment, gene delivery vehicles may be generated which carry a gene for phosphorylation, phosphoribosylation, ribosylation, or other metabolism of a purine- or pyrimidine-based drug. Such genes may have no equivalent in mammalian cells, and might come from organisms such as a virus, bacterium, fungus, or protozoan. Representative examples include: *E. coli* guanine phosphoribosyl transferase ("gpt") gene product, which converts thioxanthine into thioxanthine monophosphate (see Besnard et al., *Mol. Cell. Biol. 7*:4139-4141, 1987); alkaline phosphatase, which will convert inactive phosphorylated compounds such as mitomycin phosphate and doxorubicin-phosphate to toxic dephosphorylated compounds; fungal (*e.g.*, *Fusarium oxysporum*) or bacterial cytosine deaminase which will convert 5-fluorocytosine to the toxic compound 5-fluorouracil (Mullen, *PNAS 89*:33, 1992); carboxypeptidase G2 which will cleave the glutamic acid from para-N-bis (2-chloroethyl) aminobenzoyl glutamic acid, thereby creating a toxic benzoic acid mustard; and Penicillin-V amidase, which will convert phenoxyacetabide derivatives of doxorubicin and melphalan to toxic compounds. Conditionally lethal gene products of this type have application to many presently known purine- or pyrimidine-based anticancer drugs, which often require intracellular ribosylation or phosphorylation in order to become effective cytotoxic agents. The conditionally lethal gene product could also metabolize a nontoxic drug, which is not a purine or pyrimidine analogue, to a cytotoxic form (see Searle et al., *Brit. J. Cancer 53*:377-384, 1986).

Additionally, when the target pathogen is a mammalian virus, vector constructs may be constructed to take advantage of the fact that mammalian viruses in general tend to have "immediate early" genes, which are necessary for subsequent transcriptional activation of other viral promoter elements. Gene products of this nature are excellent candidates for intracellular signals (or "identifying agents") of viral infection. Thus, conditionally lethal genes transcribed from transcriptional promoter elements that are responsive to such viral "immediate early" gene products could specifically kill cells infected with any particular virus. Additionally, since the human a and b interferon promoter elements are transcriptionally activated in response to infection by a wide variety of nonrelated viruses, the introduction of vectors expressing a conditionally lethal gene product like HSVTK, for example, from these viral-responsive elements (VREs) could result in the destruction of cells infected with a variety of different viruses.

In another embodiment of the invention, vector constructs for incorporation into a gene delivery vehicle are provided that produce substances such as inhibitor palliatives, that inhibit viral assembly. In this context, the recombinant DNA molecule codes for defective *gag*, *pol*, *env* or other viral particle proteins or peptides which inhibit in a dominant fashion the assembly of viral particles. Such inhibition occurs because the interaction of normal subunits of the viral particle is disturbed by interaction with the defective subunits.

One way of increasing the effectiveness of inhibitory palliatives is to express inhibitory genes, such as viral inhibitory genes, in conjunction with the expression of genes which increase the probability of infection of the resistant cell by the virus in question. The result is a nonproductive "dead-end" event which would compete for productive infection events. In the specific case of HIV, a vector construct may be administered that inhibits HIV replication (by expressing anti-sense *tat*, *etc.*, as described above) and also overexpress proteins required for infection, such as CD4. In this way, a relatively small number of vector-infected HIV-resistant cells act as a "sink" or "magnet" for multiple nonproductive fusion events with free virus or virally infected cells.

In another embodiment of the invention, vector constructs are provided for the expression substances such as inhibiting peptides or proteins specific for viral protease. Viral protease cleaves the viral *gag* and *gag/pol* proteins into a number of smaller peptides. Failure of this cleavage in all cases leads to complete inhibition of production of infectious retroviral particles. The HIV protease is known to be an aspartyl protease, and these are known to be inhibited by peptides made from amino acids from protein or analogues. Vector constructs that inhibit HIV will express one or multiple fused copies of such peptide inhibitors.

Administration of the vector constructs discussed above in an appropriate gene delivery vehicle should be effective against many virally linked diseases, cancers, or other pathogenic agents.

Within still other embodiments of the invention, vector constructs are provided that code for the expression of a palliative, wherein the palliative has a membrane anchor and acts as an anti-tumor agent(s). Such a palliative may be constructed, for example, as an anti-tumor agent - membrane anchor fusion protein. Briefly, the membrane anchor aspect of the fusion protein may be selected from a variety of sequences, including, for example, the transmembrane domain of well known molecules. Generally, membrane anchor sequences are regions of a protein that bind the protein to a membrane. Customarily, there are two types of anchor sequences that attach a protein to the outer surface of a cell membrane: (1) transmembrane regions that span the lipid bilayer of the cell membrane, and interact with the hydrophobic center region

(proteins containing such regions are referred to integral membrane proteins), and (2) domains which interact with an integral membrane protein or with the polar surface of the membrane (such proteins are referred to as peripheral, or extrinsic, proteins). However, a protein can also be covalently anchored directly to a lipid component of the cell membrane.

Membrane anchors for use within the present invention may contain transmembrane domains which span the membrane one or more times. For example, in glycophorin and guanylyl cyclase, the membrane binding region spans the membrane once, whereas the transmembrane domain of rhodopsin spans the membrane seven times, and that of the photosynthetic reaction center of Rhodopseudomonas viridis spans the membrane eleven times (*see* Ross et al., *J. Biol. Chem. 257*:4152, 1982; Garbers, *Pharmac. Ther. 50*:337-345, 1991; Engelman et al., *Proc. Natl. Acad. Sci. USA 77*:2023, 1980; Heijne and Manoil, *Prot. Eng. 4*:109-112, 1990). Regardless of the number of times the protein crosses the membrane, the membrane spanning regions typically have a similar structure. More specifically, the 20 to 25 amino-acid residue portion of the domain that is located inside the membrane generally consists almost entirely of hydrophobic residues (*see* Eisenberg et al., *Ann. Rev. Biochem. 53*:595-623, 1984). For example, 28 of the 34 residues in the membrane spanning region of glycophorin are hydrophobic (*see* Ross et al., *supra*; Tomita et al., *Biochemistry 17*:4756-4770, 1978). In addition, although structures such as beta sheets and barrels do occur, the membrane spanning regions typically have an alpha helical structure, as determined by X-ray diffraction, crystallography and cross-linking studies (*see* Eisenberg et al.,*supra*; Heijne and Manoil, *supra*). The location of these transmembrane helices within a given sequence can often be predicted based on hydrophobicity plots. Stryer et al., *Biochemistry*, 3rd. ed. 304, 1988. Particularly preferred membrane anchors for use within the present invention include naturally occurring cellular proteins (that are non-immunogenic) which have been demonstrated to function as membrane signal anchors (such as glycophorin).

Within a preferred embodiment of the present invention, a vector construct is provided which encodes a membrane anchor - gamma interferon fusion protein. Within one embodiment, this fusion protein may be constructed by genetically fusing the sequence which encodes the membrane anchor of the gamma-chain of the Fc receptor, to a sequence which encodes gamma-interferon.

In yet another embodiment, vector constructs are provided which have a therapeutic effect by encoding one or more ribozymes (RNA enzymes) (Haseloff and Gerlach, *Nature 334*:585, 1989) which will cleave, and hence inactivate, RNA molecules corresponding to a pathogenic function. Since ribozymes function by recognizing a specific sequence in the target RNA and this sequence is normally 12 to 17 bp, this allows

specific recognition of a particular RNA sequence corresponding to a pathogenic state, such as HIV tat, and toxicity is specific to such pathogenic state. Additional specificity may be achieved in some cases by making this a conditional toxic palliative, as discussed above.

In still another embodiment, vector constructs are provided comprising a biologically active nucleic acid molecule that is an antisense sequence (an antisense sequence may also be encoded by a nucleic acid sequence and then produced within a host cell via transcription). In preferred embodiments, the antisense sequence is selected from the group consisting of sequences which encode influenza virus, HIV, HSV, HPV, CMV, and HBV. The antisense sequence may also be an antisense RNA complementary to RNA sequences necessary for pathogenicity. Alternatively, the biologically active nucleic acid molecule may be a sense RNA (or DNA) complementary to RNA sequences necessary for pathogenicity.

More particularly, the biologically active nucleic acid molecule may be an antisense sequence. Briefly, antisense sequences are designed to bind to RNA transcripts, and thereby prevent cellular synthesis of a particular protein, or prevent use of that RNA sequence by the cell. Representative examples of such sequences include antisense thymidine kinase, antisense dihydrofolate reductase (Maher and Dolnick, *Arch. Biochem. & Biophys. 253*:214-220, 1987; Bzik et al., *PNAS 84*:8360-8364, 1987), antisense HER2 (Coussens et al., *Science 230*:1132-1139, 1985), antisense ABL (Fainstein et al., *Oncogene 4*:1477-1481, 1989), antisense Myc (Stanton et al., *Nature 310*:423-425, 1984) and antisense *ras*, as well as antisense sequences which block any of the enzymes in the nucleotide biosynthetic pathway.

In addition, within a further embodiment of the invention antisense RNA may be utilized as an anti-tumor agent in order to induce a potent Class I restricted response. Briefly, in addition to binding RNA and thereby preventing translation of a specific mRNA, high levels of specific antisense sequences are believed to induce the increased expression of interferons (including gamma-interferon), due to the formation of large quantities of double-stranded RNA. The increased expression of gamma interferon, in turn, boosts the expression of MHC Class I antigens. Preferred antisense sequences for use in this regard include actin RNA, myosin RNA, and histone RNA. Antisense RNA which forms a mismatch with actin RNA is particularly preferred.

In another embodiment, the substances of the invention include a surface protein that is itself therapeutically beneficial. For example, in the particular case of HIV, expression of the human CD4 protein specifically in HIV-infected cells may be beneficial in two ways:

1. Binding of CD4 to HIV env intracellularly could inhibit the formation of viable viral particles much as soluble CD4 has been shown to do for free virus, but without the problem of systematic clearance and possible immunogenicity, since the protein will remain membrane bound and is structurally identical to endogenous CD4 (to which the patient should be immunologically tolerant).

2. Since the CD4/HIV env complex has been implicated as a cause of cell death, additional expression of CD4 (in the presence of excess HIV-env present in HIV-infected cells) leads to more rapid cell death and thus inhibits viral dissemination. This may be particularly applicable to monocytes and macrophages, which act as a reservoir for virus production as a result of their relative refractility to HIV-induced cytotoxicity (which, in turn, is apparently due to the relative lack of CD4 on their cell surfaces).

Still further embodiments of the present invention relate to vector constructs capable of immunostimulation. Briefly, the ability to recognize and defend against foreign pathogens is essential to the function of the immune system. In particular, the immune system must be capable of distinguishing "self" from "nonself" (*i.e.*, foreign), so that the defensive mechanisms of the host are directed toward invading entities instead of against host tissues. Cytolytic T lymphocytes (CTLs) are typically induced, or stimulated, by the display of a cell surface recognition structure, such as a processed, pathogen-specific peptide, in conjunction with a MHC class I or class II cell surface protein.

Diseases suitable to treatment by immunostimulation include viral infections such as influenza virus, respiratory syncytial virus, HPV, HBV, HCV, EBV, HIV, HSV, FeLV, FIV, Hantavirus, HTLV I, HTLV II and CMV, cancers such as melanomas, renal carcinoma, breast cancer, ovarian cancer and other cancers, and heart disease.

In one example of this embodiment, the invention provides methods for stimulating a specific immune response and inhibiting viral spread by using vector constructs that direct the expression of an antigen or modified form thereof in susceptible target cells, wherein the antigen is capable of either (1) initiating an immune response to the viral antigen or (2) preventing the viral spread by occupying cellular receptors required for viral interactions. Expression of the protein may be transient or stable with time. Where an immune response is to be stimulated to a pathogenic antigen, the vector construct is preferably designed to express a modified form of the antigen which will stimulate an immune response and which has reduced pathogenicity relative to the native antigen. This immune response is achieved when cells present antigens in the correct

manner, *i.e.*, in the context of the MHC class I and/or II molecules along with accessory molecules such as CD3, ICAM-1, ICAM-2, LFA-1, or analogs thereof (*e.g.*, Altmann et al., *Nature 338*:512, 1989).

An immune response can also be achieved by transferring to an appropriate immune cell (such as a T lymphocyte) (a) the gene for the specific T-cell receptor that recognizes the antigen of interest (in the context of an appropriate MHC molecule if necessary), (b) the gene for an immunoglobulin which recognizes the antigen of interest, or (c) the gene for a hybrid of the two which provides a CTL response in the absence of the MHC context. Thus, vector constructs may also be used as an immunostimulant, immunomodulator, or vaccine, *etc.*

In the particular case of disease caused by HIV infection, where immunostimulation is desired, the antigen generated from a vector construct may be in a form which will elicit either or both an HLA class I- or class II-restricted immune response. In the case of HIV envelope antigen, for example, the antigen is preferably selected from gp 160, gp 120, and gp 41, which have been modified to reduce their pathogenicity. In particular, the selected antigen is modified to reduce the possibility of syncytia, to avoid expression of epitopes leading to a disease enhancing immune response, to remove immunodominant, but haplotype-specific epitopes or to present several haplotype-specific epitopes, and allow a response capable of eliminating cells infected with most or all strains of HIV. The haplotype-specific epitopes can be further selected to promote the stimulation of an immune response within an animal which is cross-reactive against other strains of HIV. Antigens from other HIV genes or combinations of genes, such as *gag*, *pol*, *rev*, *vif*, *nef*, *prot*, *gag/pol*, *gag* prot, *etc.*, may also provide protection in particular cases.

HIV is only one example. This approach should be effective against many virally associated diseases or cancers where a characteristic antigen (which does not need to be a membrane protein) is expressed, such as in HPV and cervical carcinoma, HTLV-I-induced leukemias, prostate-specific antigen (PSA) and prostate cancer, mutated p53 and colon carcinoma and melanoma, melanoma specific antigens (MAGEs), and melanoma, mucin and breast cancer.

In accordance with the immunostimulation aspects of the invention, substances which are carried and/or expressed by the vector constructs of the present invention may also include "immunomodulatory factors," many of which are set forth above. Immunomodulatory factors refer to factors that, when manufactured by one or more of the cells involved in an immune response, or, which when added exogenously to the cells, causes the immune response to be different in quality or potency from that which would have occurred in the absence of the factor. The factor may also be expressed from

a gene delivery vehicle derived gene, but the expression is driven or controlled by the recombinant retrovirus. The quality or potency of a response may be measured by a variety of assays known to one of skill in the art including, for example, *in vitro* assays which measure cellular proliferation (*e.g.*, ^{3}H thymidine uptake), and *in vitro* cytotoxic assays (*e.g.*, which measure ^{51}Cr release) (*see*, Warner et al., *AIDS Res. and Human Retroviruses 7*:645-655, 1991). Immunomodulatory factors may be active both *in vivo* and *ex vivo*.

Representative examples of such factors include cytokines, such as IL-1, IL-2 (Karupiah et al., *J. Immunology 144*:290-298, 1990; Weber et al., *J. Exp. Med. 166*:1716-1733, 1987; Gansbacher et al., *J. Exp. Med. 172*:1217-1224, 1990; U.S. Patent No. 4,738,927), IL-3, IL-4 (Tepper et al., *Cell 57*:503-512, 1989; Golumbek et al., *Science 254*:713-716, 1991; U.S. Patent No. 5,017,691), IL-5, IL-6 (Brakenhof et al., *J. Immunol. 139*:4116-4121, 1987; WO 90/06370), IL-7 (U.S. Patent No. 4,965,195) , IL-8, IL-9, IL-10, IL-11, IL-12, IL-13 (*Cytokine Bulletin*, Summer 1994), IL-14 and IL-15, particularly IL-2, IL-4, IL-6, IL-12, and IL-13, alpha interferon (Finter et al., *Drugs 42*(5):749-765, 1991; U.S. Patent No. 4,892,743; U.S. Patent No. 4,966,843; WO 85/02862; Nagata et al., *Nature* 284:316-320, 1980; Familletti et al., *Methods in Enz. 78*:387-394, 1981; Twu et al., *Proc. Natl. Acad. Sci. USA 86*:2046-2050, 1989; Faktor et al., *Oncogene 5*:867-872, 1990), beta interferon (Seif et al., *J. Virol. 65*:664-671, 1991), gamma interferons (Radford et al., *The American Society of Hepatology* 20082015, 1991; Watanabe et al., *PNAS 86*:9456-9460, 1989; Gansbacher et al., *Cancer Research 50*:7820-7825, 1990; Maio et al., *Can. Immunol. Immunother. 30*:34-42, 1989; U.S. Patent No. 4,762,791; U.S. Patent No. 4,727,138), G-CSF (U.S. Patent Nos. 4,999,291 and 4,810,643), GM-CSF (WO 85/04188), tumor necrosis factors (TNFs) (Jayaraman et al., *J. Immunology 144*:942-951, 1990), CD3 (Krissanen et al., *Immunogenetics 26*:258-266, 1987), ICAM-1 (Altman et al., *Nature 338*:512-514, 1989; Simmons et al., *Nature 331*:624-627, 1988), ICAM-2, LFA-1, LFA-3 (Wallner et al., *J. Exp. Med. 166*(4):923-932, 1987), MHC class I molecules, MHC class II molecules, B7.1-.3, b_2-microglobulin (Parnes et al., *PNAS 78*:2253-2257, 1981), chaperones such as calnexin, MHC linked transporter proteins or analogs thereof (Powis et al., *Nature 354*:528-531, 1991). Within one preferred embodiment, the gene encodes gamma-interferon. Immunomodulatory factors may also be agonists, antagonists, or ligands for these molecules. For example soluble forms of receptors can often behave as antagonists for these types of factors, as can mutated forms of the factors themselves.

An example of an immunomodulatory factor cited above is a member of the B7 family of molecules (*e.g.*, B7.1-.3 costimulatory factor). Briefly, activation of the full functional activity of T cells requires two signals. One signal is provided by

interaction of the antigen-specific T cell receptor with peptides which are bound to major histocompatibility complex (MHC) molecules, and the second signal, referred to as costimulation, is delivered to the T cell by antigen presenting cells. The second signal is required for interleukin-2 (IL-2) production by T cells, and appears to involve interaction of the B7.1-.3 molecule on antigen-presenting cells with CD28 and CTLA-4 receptors on T lymphocytes (Linsley et al., *J. Exp. Med.*, *173*:721-730, 1991a and *J. Exp. Med.*, *174*:561-570, 1991). Within one embodiment of the invention, B7.1-.3 may be introduced into tumor cells in order to cause costimulation of $CD8^+$ T cells, such that the $CD8^+$T cells produce enough IL-2 to expand and become fully activated. These $CD8^+$ T cells can kill tumor cells that are not expressing B7 because costimulation is no longer required for further CTL function. Vectors that express both the costimulatory B7.1-.3 factor, and, for example, an immunogenic HBV core protein, may be made utilizing methods which are described herein. Cells transduced with these vectors will become more effective antigen presenting cells. The HBV core-specific CTL response will be augmented from the fully activated $CD8^+$ T cell via the costimulatory ligand B7.1-.3.

The choice of which immunomodulatory factor to use may be based upon known therapeutic effects of the factor, or, experimentally determined. For example, a known therapeutic effector in chronic hepatitis B infections is alpha interferon. This has been found to be efficacious in compensating a patient's immunological deficit, and thereby assisting recovery from the disease. Alternatively, a suitable immunomodulatory factor may be experimentally determined. Briefly, blood samples are first taken from patients with a hepatic disease. Peripheral blood lymphocytes (PBLs) are restimulated *in vitro* with autologous or HLA matched cells (*e.g.*, EBV transformed cells) that have been transduced with a recombinant retrovirus which directs the expression of an immunogenic portion of a hepatitis antigen and the immunomodulatory factor. These stimulated PBLs are then used as effectors in a CTL assay with the HLA matched transduced cells as targets. An increase in CTL response over that seen in the same assay performed using HLA matched stimulator and target cells transduced with a vector encoding the antigen alone, indicates a useful immunomodulatory factor. Within one embodiment of the invention, the immunomodulatory factor gamma interferon is particularly preferred.

The present invention also includes vector constructs which encode immunogenic portions of desired antigens including, for example, viral, bacterial or parasite antigens. For example, various immunogenic portions of the HBV S antigens may be combined in order to present an immune response when administered by one of the recombinant retroviruses described herein. In addition, due to the large immunological variability that is found in different geographic regions for the S antigen

open reading frame of HBV, particular combinations of antigens may be preferred for administration in particular geographic regions. Briefly, epitopes that are found in all human hepatitis B virus S samples are defined as determinant "*a*". Mutually exclusive subtype determinants however have also been identified by two-dimensional double immunodiffusion (Ouchterlony, *Progr. Allergy 5*:1, 1958). These determinants have been designated "*d*" or "*y*" and "*w*" or "*r*" (LeBouvier, *J. Infect. 123*:671, 1971; Bancroft et al., *J. Immunol. 109*:842, 1972; Courouce et al., *Bibl. Haematol. 42*:1-158, 1976). The immunological variability is due to single nucleotide substitutions in two areas of the hepatitis B virus S antigen open reading frame resulting in the following amino acid changes: (1) exchange of lysine-122 to arginine in the hepatitis B virus S antigen open reading frame causes a subtype shift from *d* to *y*, and (2) exchange of arginine-160 to lysine causes the shift from subtype *r* to *w*. In black Africa, subtype *ayw* is predominant, whereas in the U.S. and northern Europe the subtype adw_2 is more abundant (*Molecular Biology of the Hepatitis B Virus*, McLachlan (ed.), CRC Press, 1991). As will be evident to one of ordinary skill in the art, it is generally preferred to construct a vector for administration which is appropriate to the particular hepatitis B virus subtype which is prevalent in the geographical region of administration. Subtypes of a particular region may be determined by two-dimensional double immunodiffusion or, preferably, by sequencing the S antigen open reading frame of HBV virus isolated from individuals within that region.

Also presented by HBV are pol ("HBV pol"), ORF 5, and ORF 6 antigens. Briefly, the polymerase open reading frame of HBV encodes reverse transcriptase activity found in virions and core-like particles in infected liver tissue. The polymerase protein consists of at least two domains: the amino terminal domain encodes the protein that primes reverse transcription, and the carboxyl terminal domain which encodes reverse transcriptase and RNase H activity. Immunogenic portions of HBV pol may be administered to a warm-blooded animal by introducing into the animal a recombinant retrovirus which expresses the antigen of interest in order to generate an immune response within the animal. Similarly, other HBV antigens such as ORF 5 and ORF 6, (Miller et al., *Hepatology 9*:322-327, 1989), may be expressed utilizing vector constructs as described herein.

As noted above, at least one immunogenic portion of a hepatitis B antigen can be incorporated into a vector construct. The immunogenic portion(s) which are incorporated into the vector construct may be of varying length, although it is generally preferred that the portions be at least 9 amino acids long, and may include the entire antigen. Immunogenicity of a particular sequence is often difficult to predict, although T cell epitopes may be predicted utilizing the HLA A2.1 motif described by Falk et al.

(*Nature 351*:290, 1991). From this analysis, peptides may be synthesized and used as targets in an *in vitro* cytotoxic assay. Other assays, however, may also be utilized, including, for example, ELISA which detects the presence of antibodies against the newly introduced vector, as well as assays which test for T helper cells, such as gamma-interferon assays, IL-2 production assays, and proliferation assays.

Within one embodiment of the present invention, at least one immunogenic portion of a hepatitis C antigen can be incorporated into a vector construct destined for position-specific integration. Preferred immunogenic portion(s) of hepatitis C may be found in the C and NS3-NS4 regions since these regions are the most conserved among various types of hepatitis C virus (Houghton et al., *Hepatology 14*:381-388, 1991). Particularly preferred immunogenic portions may be determined by a variety of methods. For example, as noted above for the hepatitis B virus, identification of immunogenic portions of the polypeptide may be predicted based upon amino acid sequence. Briefly, various computer programs which are known to those of ordinary skill in the art may be utilized to predict CTL epitopes. For example, CTL epitopes for the HLA A2.1 haplotype may be predicted utilizing the HLA A2.1 motif described by Falk et al. (*Nature 351:290, 1991)*. From this analysis, peptides are synthesized and used as targets in an *in vitro* cytotoxic assay.

Within another aspect of the present invention, methods are provided for destroying hepatitis B carcinoma cells comprising the step of administering to a warm-blooded animal a gene delivery vehicle comprising a vector construct which directs the expression of an immunogenic portion of antigen X, such that an immune response is generated. Sequences which encode the HBxAg may readily be obtained by one of skill in the art given the disclosure provided herein. Briefly, within one embodiment of the present invention, a 642 bp Nco I-Taq I is recovered from ATCC 45020, and inserted into recombinant retroviruses as described above for other hepatitis B antigens.

The X antigen, however, is a known transactivator which may function in a manner similar to other potential oncogenes (*e.g.*, E1A). Thus, it is generally preferable to first alter the X antigen such that the gene product is non-tumorigenic before inserting it into a recombinant retrovirus. Various methods may be utilized to render the X antigen non-tumorigenic including, for example, by truncation, point mutation, addition of premature stop codons, or phosphorylation site alteration. Within one embodiment, the sequence or gene of interest which encodes the X antigen is truncated. Truncation may produce a variety of fragments, although it is generally preferable to retain greater than or equal to 50% of the encoding gene sequence. In addition, it is necessary that any truncation leave intact some of the immunogenic sequences of the gene product. Alternatively, within another embodiment of the invention, multiple translational

termination codons may be introduced into the gene. Insertion of termination codons prematurely terminates protein expression, thus preventing expression of the transforming portion of the protein.

The X gene or modified versions thereof may be tested for tumorigenicity in a variety of ways. Representative assays include tumor formation in nude mice, colony formation in soft agar, and preparation of transgenic animals, such as transgenic mice.

Within another aspect of the present invention, methods are provided for destroying hepatitis C carcinoma cells comprising the step of administering to a warm-blooded animal a recombinant retrovirus which directs the expression of an immunogenic portion of a hepatitis C antigen. Preferred immunogenic portion(s) of a hepatitis C antigen may be found in the polypeptide which contains the Core antigen and the NS1-NS5 regions (Choo et al., *Proc. Natl. Acad. Sci. USA 88*:2451-2455, 1991). Particularly preferred immunogenic portions may be determined by a variety of methods. For example, as noted above preferred immunogenic portions may be predicted based upon amino acid sequence. Briefly, various computer programs which are known to those of ordinary skill in the art may be utilized to predict CTL epitopes. For example, CTL epitopes for the HLA A2.1 haplotype may be predicted utilizing the HLA A2.1 motif described by Falk et al. (*Nature 351*:290, 1991). Another method that may also be utilized to predict immunogenic portions is to determine which portion has the property of CTL induction in mice utilizing retroviruses (*see*, Warner et al., *AIDS Res. and Human Retroviruses 7*:645-655, 1991). As noted within Warner et al., CTL induction in mice may be utilized to predict cellular immunogenicity in humans. Preferred immunogenic portions may also be deduced by determining which fragments of the polypeptide antigen or peptides are capable of inducing lysis by autologous patient lymphocytes of target cells (*e.g.*, autologous EBV-transformed lymphocytes) expressing the fragments after vector transduction of the corresponding genes.

Preferred immunogenic portions may also be selected in the following manner. Briefly, blood samples from a patient with a target disease, such as HCV, are analyzed with antibodies to individual HCV polypeptide regions (*e.g.*, HCV core, E1, E2/SNI and NS2-NS5 regions), in order to determine which antigenic fragments are present in the patient's serum. In patients treated with alpha interferon to give temporary remission, some antigenic determinants will disappear and be supplanted by endogenous antibodies to the antigen. Such antigens are useful as immunogenic portions within the context of the present invention (Hayata et al., *Hepatology 13*:1022-1028, 1991; Davis et al., *N. Eng. J. Med. 321*:1501-1506, 1989).

Additional immunogenic portions of a chosen antigen, such as those from the hepatitis B or C virus, may be obtained by truncating the coding sequence. For

example, with HBV the following sites may be truncated: Bst UI, Ssp I, Ppu M1, and Msp I (Valenzuela et al., *Nature 280*:815-19, 1979; Valenzuela et al., *Animal Virus Genetics: ICN/UCLA Symp. Mol. Cell Biol.*, 1980, B. N. Fields and R. Jaenisch (eds.), pp. 57-70, New York: Academic). Further methods for determining suitable immunogenic portions as well as methods are also described below in the context of hepatitis C.

With respect to the treatment of HBV, particularly preferred immunogenic portions for incorporation into recombinant retroviruses include HBeAg, HBcAg, and HBsAgs. Further, more than one immunogenic portion (as well as immunomodulatory factors, if desired) may be incorporated into the vector construct. For example, within one embodiment a vector construct may be prepared which directs the co-expression of both an immunogenic portion of the hepatitis B antigen, as well as an immunogenic portion of the hepatitis C polypeptide. Such constructs may be administered in order to prevent or treat acute and chronic hepatitis infections of either type B or C. Similarly, within other embodiments, a vector construct may be prepared which directs the co-expression of both an immunogenic portion of the hepatitis B X antigen, as well as an immunogenic portion of the hepatitis C polypeptide. Such a construct may similarly be administered in order to treat hepatocellular carcinoma that is associated with either hepatitis B or C. In addition, because those individuals chronically infected with hepatitis B and C are at higher risk for developing hepatocellular carcinoma, such a vector may also be utilized as a prophylactic treatment for the disease.

Immunogenic portions may also be selected by other methods. For example, the HLA A2.1/K^b transgenic mouse has been shown to be useful as a model for human T-cell recognition of viral antigens. Briefly, in the influenza and hepatitis B viral systems, the murine T-cell receptor repertoire recognizes the same antigenic determinants recognized by human T-cells. In both systems, the CTL response generated in the HLA A2.1/K^b transgenic mouse is directed toward virtually the same epitope as those recognized by human CTLs of the HLA A2.1 haplotype (Vitiello et al., *J. Exp. Med. 173*:1007-1015, 1991; Vitiello et al., *Abstract of Molecular Biology of Hepatitis B Virus Symposia*, 1992).

Immunogenic proteins of the present invention may also be manipulated by a variety of methods known in the art, in order to render them more immunogenic. Representative examples of such methods include: adding amino acid sequences that correspond to T helper epitopes; promoting cellular uptake by adding hydrophobic residues; by forming particulate structures; or any combination of these (*see generally*, Hart, op. cit., Milich et al., *Proc. Natl. Acad. Sci. USA 85*:1610-1614, 1988; Willis, *Nature 340*:323-324, 1989; Griffiths et al., *J. Virol. 65*:450-456, 1991).

The present invention also includes compositions and methods for treating, as well as vaccines for preventing, various feline diseases, including for example feline leukemia virus ("FeLV") and feline immunodeficiency virus ("FIV") infections. These viruses are discussed more fully in PCT application number US 93/09070.

Briefly, feline leukemia virus (FeLV) is a retrovirus of the oncornavirus subfamily. FeLV is presently believed to exist in three subgroups - A, B or C - which are differentiated by their envelope antigens gp70 and p15E. FeLV is also comprised of a number of core antigens, including p15, p12, p27, and p10, which are highly conserved for all subgroups of FeLV (*see* Geering et al., *Vir. 36*:678-680, 1968; Hardy et al., *JAVMA 158*:1060-1069, 1971; Hardy et al., *Science 166*:1019-1021, 1969). Within one embodiment of the invention, the vector constructs directs the expression of at least one portion of a feline leukemia virus antigen selected from the group consisting of p15*gag*, p12*gag*, p27*gag*, p10*gag*, p14*pol*, p80*pol*, p46*pol*, gp70*env*, and p15*env*. Within a particularly preferred embodiment, the vector construct directs the expression of gp85*env*. Sequences which encode these antigens may be readily obtained given the disclosure provided herein (*see* Donahue et al., *J. Vir. 62*(3):722-731, 1988; Stewart et al., *J. Vir. 58*(3):825-834, 1986; Kumar et al., *J. Vir. 63*(5):2379-2384, 1989; Elder et al., *J. Vir. 46*(3):871-880, 1983; Berry et al., *J. Vir. 62*(10):3631-3641, 1988; Laprevotte et al., *J. Vir. 50*(3):884-894, 1984).

Feline immunodeficiency virus (FIV) has been classified as a retrovirus of the lentivirus subfamily, based upon the magnesium requirement for reverse transcriptase (RT) and the morphology of viral particles (*see* Pederesen et al., *Science 235*:790-793, 1987). The feline immunodeficiency virus is morphologically and antigenically distinct from other feline retroviruses, including feline leukemia virus, type C oncorna virus (RD-114), and feline syncytium-forming virus (FeSFV) (*see* Yamamoto et al., "Efficacy of experimental FIV vaccines, (Abstract), First International Conference of Feline Immunodeficiency Virus Researchers, University of California, Davis, CA, Sep. 4-7, 1991). Within one embodiment of the invention, the vector construct directs the expression of at least one immunogenic portion of an feline immunodeficiency virus antigen selected from the group consisting of p15*gag*, p24*gag*, p10*gag*, p13*pol*, p62*pol*, p15*pol* and p36*pol*. Within a particularly preferred embodiment, the vector construct directs the expression of gp68*env*, gp27*env* and *rev*. Within the context of the present invention, "*rev*" is understood to refer to the antigen corresponding to the *rev* open reading frame (*see*, Phillips et al., First International Conference, *supra*). Sequences which encode these antigens may be readily obtained by one of skill in the art given the disclosure provided herein (*see* Phillips et al., *J. Vir. 64*(10):4605-4613, 1990; Olmsted et al., *PNAS 86*:2448-2452, 1989; Talbott et al., *PNAS 86*:5743-5747, 1989).

Still other examples include vector constructs which direct the expression of a non-tumorigenic, altered genes such as the ras (ras*) gene (*see* WO 93/10814; U.S.S.N. 08/367,071, *supra*), an altered p53 (p53*) gene (Linzer and Levine, *Cell 17*:43-52, 1979; Lane and Crawford, *Nature 278*:261-263, 1979; Hinds et al., *J. Virol. 63*:739-746, 1989; U.S.S.N. 08/367,071, *supra*), an altered Rb (Rb*) gene (Friend et al., *Nature 323*:643, 1986; Lee et al., *Science 235*:1394, 1987; Fung et al., *Science 236*:1657, 1987; U.S.S.N. 08/367,071, *supra*), an altered gene which causes Wilms' tumor (Call et al., *Cell 60*:509, 1990; Gessler et al., *Nature 343*:744, 1990; Rose et al., *Cell 60*:495, 1990; Haber et al., *Cell 61*:1257, 1990; U.S.S.N. 08/367,071, *supra*), an altered mucin gene (Girling et al., *Int. J. Cancer 43*:1072-1076, 1989; Gendler et al., *J. Biol. Chem. 265*(25):15286-15293, 1990; Lan et al., *J. Biol. Chem. 265*(25):15294-15299, 1990; Ligtenberg et al., *J. Biol. Chem. 265*:5573-5578, 1990; Jerome et al., *Cancer Res. 51*:2908-2916, 1991; and U.S.S.N. 08/367,071, *supra*), an altered DCC (deleted in colorectal carcinomas) gene (reviewed by Edelman in *Biochem 27*:3533-3543, 1988; Solomon, *Nature 343*:412-414, 1990; Fearon et al., *Science 247*:49-56, 1990; U.S.S.N. 08/367,071, *supra*), a MCC (mutated in colorectal cancer) or APC. Both MCC and APC gene (Kinzler et al., *Science 251*:1366-1370, 1991; Nishiho et al., *Science 253*:665-669, 1991), an altered cellular receptor, including, for example, neu and mutated or altered forms of the thyroid hormone receptor, the PDGF receptor, the insulin receptor, the Interleukin receptors (*e.g.*, IL-1, -2, -3, etc. receptors), or the CSF receptors, such as the G-CSF, GM-CSF, or M-CSF receptors (U.S.S.N. 08/367,071, *supra*, Slamon et al., *Science 244*:707-712, 1989; Slamon et al., *Cancer Cells 7*:371-380, 1989; Shih et al., *Nature 290*:261, 1981; Schechter, *Nature 312*:513, 1984; Coussens et al., *Science 230*:1132, 1985; Leduc et al., *Am. J. Hum. Genet. 44*:282-287, 1989). Alterations in such receptor result in the production of protein(s) (or receptors) containing novel coding sequence(s). The novel protein(s) encoded by these sequence(s) may be used as a marker of tumorigenic cells, and an immune response directed against these novel coding region(s) may be utilized to destroy tumorigenic cells containing the altered sequence(s) or gene(s).

If the altered cellular component is associated with making the cell tumorigenic, then, it is necessary to make the altered cellular component non-tumorigenic. For example, within one embodiment, the sequence or gene of interest which encodes the altered cellular component is truncated in order to render the gene product non-tumorigenic. The gene encoding the altered cellular component may be truncated to a variety of sizes, although it is preferable to retain as much as possible of the altered cellular component. In addition, it is necessary that any truncation leave intact at least some of the immunogenic sequences of the altered cellular component. Alternatively, multiple translational termination codons may be introduced into the gene which encodes

the altered cellular component, downstream of the immunogenic region. Insertion of termination codons will prematurely terminate protein expression, thus preventing expression of the transforming portion of the protein. *See* U.S.S.N. 08/367,071, *supra*. It should be noted, however, that if the altered cellular component is only associated with non-tumorigenic cells in general, and is not required or essential for making the cell tumorigenic, then it is not necessary to render the cellular component non-tumorigenic. Representative examples of such altered cellular components which are not tumorigenic include Rb*, ubiquitin*, and mucin*.

As noted above, in order to generate an appropriate immune response, the altered cellular component must also be immunogenic. Immunogenicity of a particular sequence is often difficult to predict, although T cell epitopes often possess an immunogenic amphipathic alpha-helix component. In general, however, it is preferable to determine immunogenicity in an assay. Representative assays include an ELISA which detects the presence of antibodies against the newly introduced vector, as well as assays which test for T helper cells such as gamma-interferon assays, IL-2 production assays, and proliferation assays. A particularly preferred method for determining immunogenicity is the CTL assay. *See* U.S.S.N. 08/367,071, *supra*),

Once a sequence encoding at least one anti-tumor agent has been obtained, it is preferable to ensure that the sequence encodes a non-tumorigenic protein. Various assays are known and may easily be accomplished which assess the tumorigenicity of a particular cellular component. Representative assays include tumor formation in nude mice or rats, colony formation in soft agar, and preparation of transgenic animals, such as transgenic mice.

For this and many other aspects of the invention, tumor formation in nude mice or rats is a particularly important and sensitive method for determining the tumorigenicity of an anti-tumor agent. Nude mice lack a functional cellular immune system (*i.e.*, do not possess CTLs), and therefore provide a useful *in vivo* model in which to test the tumorigenic potential of cells. Normal non-tumorigenic cells do not display uncontrolled growth properties if injected into nude mice. However, transformed cells will rapidly proliferate and generate tumors in nude mice. Briefly, in one embodiment the vector construct encoding the altered cellular component is delivered to syngeneic murine cells, followed by administration into nude mice. The mice are visually examined for a period of 2 to 8 weeks after administration in order to determine tumor growth. The mice may also be sacrificed and autopsied in order to determine whether tumors are present. (Giovanella et al., *J. Natl. Cancer Inst. 48*:1531-1533, 1972; Furesz et al., "Tumorigenicity testing of cell lines considered for production of biological drugs," *Abnormal Cells*, New Products and Risk, Hopps and Petricciani (eds.), Tissue Culture

Association, 1985; and Levenbook et al., *J. Biol. Std. 13*:135-141, 1985). Tumorigenicity may also be assessed by visualizing colony formation in soft agar (MacPherson and Montagnier, *Vir. 23*:291-294, 1964). Briefly, one property of normal non-tumorigenic cells is "contact inhibition" (*i.e.*, cells will stop proliferating when they touch neighboring cells). If cells are plated in a semi-solid agar support medium, normal cells rapidly become contact inhibited and stop proliferating, whereas tumorigenic cells will continue to proliferate and form colonies in soft agar.

Transgenic animals, such as transgenic mice, may also be utilized to assess the tumorigenicity of an anti-tumor agent (*e.g.*, Stewart et al., *Cell 38*:627-637, 1984; Quaife et al., *Cell 48*:1023-1034, 1987; and Koike et al., *Proc. Natl. Acad. Sci. USA 86*:5615-5619, 1989). In transgenic animals, the gene of interest may be expressed in all tissues of the animal. This unregulated expression of the transgene may serve as a model for the tumorigenic potential of the newly introduced gene.

In addition to tumorigenicity studies, it is generally preferable to determine the toxicity of a gene delivery vehicle prior to administration. A variety of methods well known to those of skill in the art may be utilized to measure such toxicity, including for example, clinical chemistry assays which measure the systemic levels of various proteins and enzymes, as well as blood cell volume and number.

The present invention also provides vector constructs encoding one or more gene products capable of immune down-regulation. Briefly, specific down-regulation of inappropriate or unwanted immune responses, such as in autoimmune or pseudo-autoimmune diseases such as chronic hepatitis, diabetes, rheumatoid arthritis, graft vs. host disease and Alzheimer's, or in transplants of heterologous tissue such as bone marrow, can be engineered using immune-suppressive viral gene products, or active portion thereof, which suppress surface expression of transplantation (MHC) antigen. Within the present invention, an "active portion" of a gene product is that fragment of the gene product which must be retained for biological activity. Such fragments or active domains can be readily identified by systematically removing nucleotide sequences from the protein sequence, transforming target cells with the resulting vector construct, and determining MHC class I presentation on the surface of cells using FACS analysis or other immunological assays, such as a CTL assay. These fragments are particularly useful when the size of the sequence encoding the entire protein exceeds the capacity of the viral carrier. Alternatively, the active domain of the MHC antigen presentation inhibitor protein can be enzymatically digested and the active portion purified by biochemical methods. For example, a monoclonal antibody that blocks the active portion of the protein can be used to isolate and purify the active portion of the cleaved protein (Harlow et al., *Antibodies: A Laboratory Manual*, Cold Springs Harbor, 1988).

For example, suppression is effected by specifically inhibiting the activation of display of processed peptides in the context of self MHC molecules along with accessory molecules such as CD8, intercellular adhesion molecule -1 (ICAM-1), ICAM-2, ICAM-3, leukocyte functional antigen-1 (LFA-1) (Altmann et al., *Nature 338*:521, 1989), the B7.1-.3 molecule (Freeman et al., *J. Immunol. 143*:2714, 1989), LFA-3 (Singer, *Science 255*:1671, 1992; Rao, *Crit. Rev. Immunol. 10*:495, 1991), or other cell adhesion molecules. Antigenic peptide presentation in association with MHC class I molecules leads to CTL activation. Transfer and stable integration of specific sequences capable of expressing products expected to inhibit MHC antigen presentation block activation of T-cells, such as $CD8^+$ CTL, and therefore suppress graft rejection. A standard CTL assay may be utilized in order to detect this response. Components of the antigen presentation pathway include the 45 Kd MHC class I heavy chain, b_2-microglobulin, processing enzymes such as proteases, accessory molecules, chaperones such as calnexin (Gaczynska, *et al.*, *Nature*, *365*: 264-282, 1993), and transporter proteins such as PSF1, TAP1 and TAP 2 (Driscoll, *et al.*, *Nature*, *365*: 262-263, 1993).

In an alternative example, vector constructs are provided which direct the expression of a gene product or an active portion of a gene product capable of binding b_2-microglobulin. Briefly, transport of MHC class I molecules to the cell surface for antigen presentation requires association with b_2-microglobulin. Thus, proteins that bind b_2-microglobulin and inhibit its association with MHC class I indirectly inhibit MHC class I antigen presentation. Suitable proteins include the H301 gene product. Briefly, the H301 gene, obtained from the human cytomegalovirus (CMV) encodes a glycoprotein with sequence homology to the b_2-microglobulin binding site on the heavy chain of the MHC class I molecule (Browne et al., *Nature 347*:770, 1990). H301 binds b_2-microglobulin, thereby preventing the maturation of MHC class I molecules, and renders transformed cells unrecognizable by cytotoxic T-cells, thus evading MHC class I restricted immune surveillance.

Within another embodiment, vector constructs are provided which direct the expression of a protein or active portion of a protein that binds to newly synthesized MHC class I molecules intracellularly. Alternatively, antisense RNA or ribozymes which inhibit translation of MHC class I proteins are encoded by the vector construct. These activities prevent migration of the MHC class I molecule from the endoplasmic reticulum, resulting in the inhibition of terminal glycosylation. This blocks transport of these molecules to the cell surface and prevents cell recognition and lysis by CTL. For instance, one of the products of the E3 gene may be used to inhibit transport of MHC class I molecules to the surface of the transformed cell. More specifically, E3 encodes a 19 kD transmembrane glycoprotein, E3/19K, transcribed from the E3 region of the adenovirus 2

genome. Within the context of the present invention, tissue cells are transformed with a vector construct containing the E3/19K sequence, which upon expression produces the E3/19K protein. The E3/19K protein inhibits the surface expression of MHC class I surface molecules, and cells transformed by the recombinant retrovirus evade an immune response. Consequently, donor cells can be transplanted with reduced risk of graft rejection and may require only a minimal immunosuppressive regimen for the transplant patient. This allows an acceptable donor-recipient chimeric state to exist with fewer complications. Similar treatments may be used to treat the range of so-called autoimmune diseases, including systemic lupus erythematosus, multiple sclerosis, rheumatoid arthritis or chronic hepatitis B infection.

Another alternative method of immunosuppression involves the use of antisense message, ribozyme, or other gene expression inhibitor specific for T-cell clones which are autoreactive in nature. These block the expression of the T-cell receptor of particular unwanted clones responsible for an autoimmune response. The antisense, ribozyme, or other gene may be introduced using a gene delivery vehicles.

Other proteins, not discussed above, that function to inhibit, suppress or down-regulate MHC class I antigen presentation (MHC class II presentation) may also be identified and utilized within the context of the present invention. In order to identify such proteins, in particular those derived from mammalian pathogens (and, in turn, active portions thereof), a recombinant retrovirus that expresses a protein or an active portion thereof suspected of being capable of inhibiting MHC class I antigen presentation is transformed into a tester cell line, such as BC. The tester cell lines with and without the sequence encoding the candidate protein are compared to stimulators and/or targets in the CTL assay. A decrease in cell lysis corresponding to the transformed tester cell indicates that the candidate protein is capable of inhibiting MHC presentation.

Many infectious diseases, cancers, autoimmune diseases, and other diseases involve the interaction of viral particles with cells, cells with cells, or cells with factors. In viral infections, viruses commonly enter cells via receptors on the surface of susceptible cells. In cancers, cells may respond inappropriately or not at all to signals from other cells or factors. In autoimmune disease, there is inappropriate recognition of "self" markers. Within the present invention, such interactions may be blocked by utilizing vector constructs that code for the *in vivo* expression of an analogue to either of the partners in an interaction. Such an analogue is known as a blocking agent. This blocking action may occur intracellularly, on the cell membrane, or extracellularly. The blocking action of a blocking agent can be mediated either from inside a susceptible cell or by secreting a version of the blocking protein to locally block the pathogenic interaction.

For example, in the case of HIV, the two agents of interaction are the gp 120/gp 41 envelope protein and the CD4 receptor molecule. Thus, an appropriate blocker would be an HIV env analogue that blocks HIV entry without causing pathogenic effects, or a CD4 receptor analogue. The CD4 analogue would be secreted and would function to protect neighboring cells, while the gp 120/gp 41 is secreted or produced only intracellularly so as to protect only the vector-containing cell. It may be advantageous to add human immunoglobulin heavy chains or other components to CD4 in order to enhance stability or complement lysis. Delivery of a vector construct encoding such a hybrid-soluble CD4 in a position-specific manner results in a continuous supply of a stable hybrid molecule.

Recombinant DNA molecules mediating the expression of HIV *env* may also be constructed. It will be evident to one skilled in the art which portions are capable of blocking virus adsorption without overt pathogenic side effects (Willey et al., *J. Virol. 62*:139, 1988; Fisher et al., *Science 233*:655, 1986).

Another embodiment of the invention involves the delivery of suppressor genes which, when deleted, mutated or not expressed in a cell type, lead to tumorigenesis in that cell type. Reintroduction of the deleted gene leads to regression of the tumor phenotype in these cells. Since malignancy can be considered to be an inhibition of cellular terminal differentiation compared with cell growth, the delivery and expression of gene products which lead to differentiation of a tumor should also, in general, lead to regression.

Sequences which encode the above-described nucleic acid molecules may be obtained from a variety of sources. For example, plasmids which contain sequences that encode altered cellular products may be obtained from a depository such as the American Type Culture Collection (ATCC, Rockville, Maryland), or from commercial sources such as Advanced Biotechnologies (Columbia, Maryland) or British Bio-Technology Limited (Cowley, Oxford England). Alternatively, cDNA sequences for use with the present invention may be obtained from cells which express or contain the sequences. Briefly, within one embodiment mRNA from a cell which expresses the gene of interest is reverse transcribed with reverse transcriptase using oligo dT or random primers. The single stranded cDNA may then be amplified by PCR (*see* U.S. Patent Nos. 4,683,202, 4,683,195 and 4,800,159. See also *PCR Technology: Principles and Applications for DNA Amplification*, Erlich (ed.), Stockton Press, 1989) utilizing oligonucleotide primers complementary to sequences on either side of desired sequences. In particular, a double stranded DNA is denatured by heating in the presence of heat stable Taq polymerase, sequence specific DNA primers, ATP, CTP, GTP and TTP. Double-stranded DNA is produced when synthesis is complete. This cycle may be

repeated many times, resulting in a factorial amplification of the desired DNA. In addition, genes of known nucleotide sequence useful for practice of the invention may be obtained by cloning the desired gene from an available cDNA or mRNA library using standard techniques. Nucleic acid molecules which are carried and/or expressed by the recombinant DN molecules described herein may also be synthesized, in whole or in part, for example, on an Applied Biosystems Inc. DNA synthesizer (*e.g.*, APB DNA synthesizer model 392 (Foster City, California).

C. CELL CULTURE

As noted above, certain embodiments of the present invention provide viral preparations suitable for administration to humans.

In a particularly preferred embodiment, high titer recombinant retroviral preparations are employed. To produce such high titer preparations, cell culture methods, as described in U.S.S.N. 08/367,071, *supra*, are used. Briefly, a wide variety of methods may be utilized, including for example, the use of fermenters or bioreactors, roller bottles, cell hotels or cell factories, and hollow fiber culture. For bioreactors or fermenters, cells are preferably grown on microcarriers (*i.e.*, Cytodex 1 or Cytodex 2; Pharmacia, Piscataway, N.J.) at concentrations ranging from 3 to 15 grams of microcarrier per liter of appropriate media. For roller bottles, suitable conditions include those used for bioreactors, with the exception that microcarrier beads are not utilized. Cell factories may also be utilized for the large scale cell culture and production of recombinant retroviruses from adherent cells. Briefly, cell factories (also termed "cell hotels") typically contain multiple trays molded from virgin polystyrene assembled by sonic welding one to another, followed by treatment to provide a Nuclon D surface, and are available from a variety of manufacturers, including for example Nunc. Hollow fiber culture methods may also be used to produce recombinant retroviruses. Briefly, high titer retroviral production using hollow fiber cultures is based on increasing viral concentration as the cells are being cultured to a high density in a reduced volume of media. The volume of media being cultured on the cell side is approximately 10 to 100 fold lower then volumes required for equivalent cell densities cultured in tissue culture dishes or flasks, resulting in a 10-100 fold increase in titer when an individual retroviral producer cell line is amenable to hollow fiber growth conditions. To achieve maximum cell density, the individual cells must be able to grow in very close proximity and on top of each other. Many cell lines will not grow in this fashion and retroviral packaging cell lines based on these types of cell lines may not achieve 10 fold increases in titer. Cell lines which would grow very well would be non-adherent cell line and it is believed that a retroviral producer line based on a non-

adherent cell line may reach 100 fold increases in titer compared to tissue culture dishes and flasks.

D. CONCENTRATION AND PURIFICATION OF RECOMBINANT RETROVIRAL PARTICLES

When a retroviral gene delivery vehicle is employed in the practice of the present invention, it is preferable to increase the purity of therapeutic preparation, as well as to increase the titer of recombinant retrovirus that may be administered. A wide variety of methods may be utilized for increasing viral concentration and purity, including for example, precipitation of recombinant retroviruses with ammonium sulfate, polyethylene glycol ("PEG") concentration, concentration by centrifugation (either with or without gradients such as PERCOLL, or "cushions" such as sucrose, use of concentration filters (*e.g.*, Amicon filtration), and 2-phase separations. *See* U.S.S.N. 08/367,071, *supra*, and U.S.S.N. 08/153,342.

E. ASSAYS

Within other aspects of the present invention, methods are provided for quantitating gene delivery vehicles utilizing non-denaturing gels (*e.g.*, 4-15% gradient polyacrylamide gels for resolving recombinant retroviral particles), along with methods for estimating or quantitating the resultant products such as, for example, staining with coomassie blue or silver stain, followed by densitometry scanning. Such methods, while not capable of discriminating between viable and non-viable gene delivery vehicles, are advantageous because they are relatively simple and quick. One representative example of such methods is set forth below in Example 10 in more detail.

Within other aspects of the present invention, assays are provided for titering recombinant virus, such as a recombinant retrovirus, in a sample. Typically, such assays may be based upon presence of a selectable marker, or formation of blue colonies. However, within certain embodiments gene delivery vehicles are provided which do not include a gene coding for a selectable marker. Therefore, antibody and PCR assays, the latter of which is described below, may be employed in order to determine titer. To use PCR to amplify sequences unique to the vector construct carried in the gene delivery vehicle, appropriate amplification primers are required. Such primers can readily be designed by those skilled in the art and will depend on the vector construct employed, the components thereof, the particular region(s) desired to be amplified, *etc.* When a recombinant retrovirus is used as the gene delivery, representative PCR primers pairs include those specific for LTR sequences, packaging signal sequences or other regions of the retroviral backbone, and may include primers specific for the gene(s) of interest.

Briefly, within one embodiment of the invention a PCR titering assay is used to determine titer of a recombinant retroviral preparation. The assay is performed by growing a known number of cells transduced with the recombinant retrovirus on 6-well plates for at least 16 hr. before harvest. One well per plate is sacrificed for counting. Cells from the other wells are lysed and their contents isolated. DNA is prepared using a QUIAmp DNA isolation kit (QUIAgen, Inc.,Chatsworth, CA). DNAs are resuspended in 5×10^6 cell equivalents/μL per sample.

To calculate titer, a standard curve is generated using DNA isolated from 5×10^6 untransduced HT1080 cells (negative control) and 5×10^6 HT1080 cells transduced with a known vector and having one integrated copy of that vector per cell genome (positive control), such as may be prepared from packaging cell lines transduced with a recombinant retrovirus encoding a selectable marker, *e.g.*, neomycin resistance. The standard curve is generated by combining different amounts of the positive and negative control DNA and amplifying specific sequences therefrom by PCR using primers specific to a particular region of the recombinant retrovirus. A representative group of mixtures for generating a standard curve is:

Tube	100%	75%	50%	25%	10%	5%	0%	Blank
Positive Control (μL)	50	37.5	25	12.5	5	2.5	0	0
Negative Control (μL)	0	12.5	25	37.5	45	47.5	50	0
Distilled water (μL)	0	0	0	0	0	0	0	50

Five microliters from each tube is placed into one of eight reaction tubes (duplicates are also prepared), with the remainder bring stored at -20_C. Five microliters from each sample DNA preparations are placed into their own reaction tubes in duplicate. PCR reactions (50 μL total volume) are then initiated by adding 45.0 μL of a reaction mix containing the following components per tube to be tested: 24.5 μL water, 5 μL 10X reaction PCR buffer, 4 μL of 25 mM $MgCl_2$, 4 μL dNTPs (containing 2.5 mM of each of dATP, dGTP, dCTP, and dTTP), 5 μL of primer mix (100 ng or each primer), 0.25 μL TaqStart monoclonal antibody (Clontech Laboratories, Inc., Palo Alto, CA), 1.00 μL TaqStart buffer (Clontech Labs, Inc.), and 0.25 μL AmpliTaq DNA polymerase (Perkin-Elmer, Inc., Norwalk, CN). Just prior to aliquoting the reaction mix to the reaction tubes, 1 μL of α-^{32}P dCTP (250 μCi; 3000 C/mmol, 10 mCi/mL, Amersham Corp., Arlington Heights, IL) is added into the reaction mix. After aliquoting 45.0 μL the reaction mix into each of the reaction tubes, the tubes are capped and placed into a thermocycler. The particular denaturation, annealing, elongation times and temperatures, and number of thermocycles will vary depending on size and nucleotide composition of the primer pair

used. 20 - 25 amplification thermocycles are then performed. 5 μL of each reaction is then spotted on DE81 ion exchange chromatography paper (Whatman, Maidstone, England) and air dried for 10 min. The filter is then washed five times, 100 mL per wash, in 50 mM Na_2PO_4, pH 7, 200 mM NaCl, after which it is air dried and then sandwiched in Saran Wrap. Quantitation is performed on a PhosphoImager SI (Molecular Dynamics, Sunnyvale, CA). Filters are typically exposed to a phosphor screen, which stores energy from ionizing radiation, for a suitable period, typically about 120 min. After exposure, the phosphor screen is scanned, whereby light is emitted in proportion to the radioactivity on the original filter. The scanning results are then downloaded and plotted on a log scale as cpm (ordinate) versus percent positive control DNA (abscissa). Titers (infectious units/mL) for each sample are calculated by multiplying the number of cells from which DNA was isolated by the percentage (converted to decimal form) determined from the standard curve based on the detected radioactivity, divided by the volume of recombinant retrovirus used to transduce the cells. As will be appreciated by those in the art, other methods of detection, such as colorimetric methods, may also be employed to label the amplified products.

F. FORMULATION

Another aspect of the invention relates to formulation of the gene delivery vehicles according to the invention.

A preferred embodiment of this aspect of the invention concerns the preservation of an infectious recombinant retroviral preparation, such that the recombinant retroviruses are capable of infecting eukaryotic cells upon reconstitution (see U.S. Serial No. 08/153,342) and delivering the vector construct into the genome of such cell in a position-specific manner. Briefly, purified, and optionally concentrated, recombinant retrovirus may be preserved by first adding a sufficient amount of a formulation buffer to the media containing the recombinant retrovirus to form an aqueous suspension. Preferably, the formulation buffer is an aqueous solution that contains a saccharide, a high molecular weight structural additive, and a buffering component in water. As utilized within the context of the present invention, a "buffering compound" or "buffering component" should be understood to refer to a substance that functions to maintain the aqueous suspension at a desired pH. The aqueous solution may also contain one or more amino acids.

The recombinant retrovirus can also be preserved in a purified form. More specifically, prior to the addition of the formulation buffer, the crude recombinant

retrovirus described above may be clarified by passing it through a filter, and then concentrated, such as by a cross flow concentrating system (Filtron Technology Corp., Nortborough, MA). Within one embodiment, DNase is added to the concentrate to digest exogenous DNA. The digest is then diafiltrated to remove excess media components and establish the recombinant retrovirus in a more desirable buffered solution. The diafiltrate is then passed over a Sephadex S-500 gel column and a purified recombinant retrovirus is eluted. A sufficient amount of formulation buffer is added to this eluate to reach a desired final concentration of the constituents (see, *e.g.*, Example 9) and to minimally dilute the recombinant retrovirus, and the aqueous suspension is then stored, preferably at -70°C or immediately dried. As noted above, the formulation buffer is an aqueous solution that contains a saccharide, a high molecular weight structural additive, and a buffering component in water. The aqueous solution may also contain one or more amino acids.

The crude recombinant retrovirus can also be purified by ion exchange column chromatography. This method is described in more detail in U.S. Patent Application Serial No. 08/093,436. In general, the crude recombinant retrovirus is clarified by passing it through a filter, and the filtrate loaded onto a column containing a highly sulfonated cellulose matrix. The recombinant retrovirus is eluted from the column in purified form by using a high salt buffer. The high salt buffer is then exchanged for a more desirable buffer by passing the eluate over a molecular exclusion column. A sufficient amount of formulation buffer is then added, as discussed above, to the purified recombinant retrovirus and the aqueous suspension is either dried immediately or stored, preferably at -70°C.

The aqueous suspension in crude or purified form can be dried by lyophilization or evaporation at ambient temperature. Specifically, lyophilization involves the steps of cooling the aqueous suspension below the glass transition temperature or below the eutectic point temperature of the aqueous suspension, and removing water from the cooled suspension by sublimation to form a lyophilized retrovirus. Briefly, aliquots of the formulated recombinant retrovirus are placed into an Edwards Refrigerated Chamber (3 shelf RC3S unit) attached to a freeze dryer (Supermodulyo 12K). A multistep freeze drying procedure as described by Phillips et al. (Cryobiology 18:414, 1981) is used to lyophilize the formulated recombinant retrovirus, preferably from a temperature of -40_C to -45_C. The resulting composition contains less than 10% water by weight of the lyophilized retrovirus. Once lyophilized, the recombinant retrovirus is stable and may be stored at -20_C to 25_C, as discussed in more detail below.

Within the evaporative method, water is removed from the aqueous suspension at ambient temperature by evaporation. Within one embodiment, water is removed through spray drying (EP 520,748). Within the spray drying process, the

aqueous suspension is delivered into a flow of preheated gas, usually air, whereupon water rapidly evaporates from droplets of the suspension. Spray drying apparatus are available from a number of manufacturers (e.g., Drytec, Ltd., Tonbridge, England; Lab-Plant, Ltd., Huddersfield, England). Once dehydrated, the recombinant retrovirus is stable and may be stored at -20_C to 25_C. Within the methods described herein, the resulting moisture content of the dried or lyophilized retrovirus may be determined through use of a Karl-Fischer apparatus (EM Science Aquastar' V1B volumetric titrator, Cherry Hill, NJ), or through a gravimetric method.

The aqueous solutions used for formulation, as previously described, are composed of a saccharide, high molecular weight structural additive, a buffering component, and water. The solution may also include one or more amino acids. The combination of these components act to preserve the activity of the recombinant retrovirus upon freezing and lyophilization, or drying through evaporation. Although a preferred saccharide is lactose, other saccharides may be used, such as sucrose, mannitol, glucose, trehalose, inositol, fructose, maltose or galactose. In addition, combinations of saccharides can be used, for example, lactose and mannitol, or sucrose and mannitol. A particularly preferred concentration of lactose is 3%-4% by weight. Preferably, the concentration of the saccharide ranges from 1% to 12% by weight.

The high molecular weight structural additive aids in preventing viral aggregation during freezing and provides structural support in the lyophilized or dried state. Within the context of the present invention, structural additives are considered to be of "high molecular weight" if they are greater than 5000 m.w. A preferred high molecular weight structural additive is human serum albumin. However, other substances may also be used, such as hydroxyethyl-cellulose, hydroxymethyl-cellulose, dextran, cellulose, gelatin, or povidone. A particularly preferred concentration of human serum albumin is 0.1% by weight. Preferably, the concentration of the high molecular weight structural additive ranges from 0.1% to 10% by weight.

The amino acids, if present, function to further preserve viral infectivity upon cooling and thawing of the aqueous suspension. In addition, amino acids function to further preserve viral infectivity during sublimation of the cooled aqueous suspension and while in the lyophilized state. A preferred amino acid is arginine, but other amino acids such as lysine, ornithine, serine, glycine, glutamine, asparagine, glutamic acid or aspartic acid can also be used. A particularly preferred arginine concentration is 0.1% by weight. Preferably, the amino acid concentration ranges from 0.1% to 10% by weight.

The buffering component acts to buffer the solution by maintaining a relatively constant pH. A variety of buffers may be used, depending on the pH range desired, preferably between 7.0 and 7.8. Suitable buffers include phosphate buffer and

citrate buffer. A particularly preferred pH of the recombinant retrovirus formulation is 7.4, and a preferred buffer is tromethamine.

In addition, it is preferable that the aqueous solution contain a neutral salt which is used to adjust the final formulated recombinant retrovirus to an appropriate iso-osmotic salt concentration. Suitable neutral salts include sodium chloride, potassium chloride or magnesium chloride. A preferred salt is sodium chloride.

Aqueous solutions containing the desired concentration of the components described above may be prepared as concentrated stock solutions.

A particularly preferred method of preserving recombinant retroviruses in a lyophilized state for subsequent reconstitution comprises the steps of (a) combining an infectious recombinant retrovirus with an aqueous solution to form an aqueous suspension, the aqueous suspension including 4% by weight of lactose, 0.1% by weight of human serum albumin, 0.03% or less by weight of NaCl, 0.1% by weight of arginine, and an amount of tromethamine buffer effective to provide a pH of the aqueous suspension of approximately 7.4, thereby stabilizing the infectious recombinant retrovirus; (b) cooling the suspension to a temperature of from -40_C to -45_C to form a frozen suspension; and (c) removing water from the frozen suspension by sublimation to form a lyophilized composition having less than 2% water by weight of the lyophilized composition, the composition being capable of infecting mammalian cells upon reconstitution. It is preferred that the recombinant retrovirus be replication defective and suitable for administration into humans upon reconstitution.

It will be evident to those skilled in the art given the disclosure provided herein that it may be preferable to utilize certain saccharides within the aqueous solution when the lyophilized retrovirus is intended for storage at room temperature. More specifically, it is preferable to utilize disaccharides, such as lactose or trehalose, particularly for storage at room temperature.

The lyophilized or dehydrated retroviruses of the subject invention may be reconstituted using a variety of substances, but are preferably reconstituted using water. In certain instances, dilute salt solutions which bring the final formulation to isotonicity may also be used. In addition, it may be advantageous to use aqueous solutions containing components known to enhance the activity of the reconstituted retrovirus. Such components include cytokines, such as IL-2, polycations, such as protamine sulfate, or other components which enhance the transduction efficiency of the reconstituted retrovirus. Lyophilized or dehydrated recombinant retrovirus may be reconstituted with any convenient volume of water or the reconstituting agents noted above that allow substantial, and preferably total solubilization of the lyophilized or dehydrated sample.

G. ADMINISTRATION

In another aspect of the present invention, methods are provided for treating eukaryotic organisms, particularly warm blooded animals, and especially human, afflicted with a variety of diseases, including a genetic disease, cancer, an infectious disease, an autoimmune disease, and inflammatory disease, a cardiovascular disease, and a degenerative disease. Examples of genetic diseases include but are not limited to; thalassemia, phenylketonuria, Lesch-Nyan syndrome, SCID, hemophilia A and B, cystic fibrosis, Duchenne's muscular dystrophy, inherited emphysema, familial hypercholesterolemia, and Gaucher's disease. Examples of cancers include but are not limited to; solid tumors, leukemias and lymphomas. Representative examples include melanomas, colorectal carcinomas, lung carcinomas (including large cell, small cell, squamous and adeno-carcinomas), renal cell carcinomas, cervical cancer, adult T-cell lymphoma leukemia, and breast adeno-carcinomas. Infectious diseases include but not limited to; hepatitis, tuberculosis, malaria, human immunodeficiency virus, herpes virus, tetanus, dysentery, shigella, FeLV, and FIV. Degenerative diseases include but are not limited to: Alzheimer's disease, multiple sclerosis, muscular dystrophy, amyotrophic lateral sclerosis, Inflammatory diseases include rheumatoid arthritis, spinal meningitis, and pancreatitis. Autoimmune diseases include diabetes, uveitis, HIV, and SCID. Cardiovascular diseases include, chronic rheumatic heart disease, arteriosclerosis, mitral valve and aortic stenosis, myocarditis, pericarditis, Marfan's syndrome, Ehlers-Danlos syndrome, Churg-Strauss syndrome, and scleroderma.

Each of these methods comprise administration of a gene delivery vehicle according to the invention such that a therapeutically efficacious amount of gene product encoded by the gene of interest carried thereby is produced. As used herein, a "therapeutically effective amount" of a gene product expressed from a vector construct according to the invention is an amount that achieves a desired therapeutic benefit in a patient to an extent greater than that observed when the patient was not treated with the gene product. For instance, when the gene product is factor VIII, a "therapeutically effective amount" refers to the amount of factor VIII needed to produce therapeutically beneficial clotting and will thus generally be determined by each patient's attending physician, although serum levels of about 0.2 ng/mL (about 0.1% of "normal" levels) or more will typically be therapeutically beneficial. When the gene product is an RNA molecule with intrinsic biological activity, such an antisense RNA or ribozyme, a "therapeutically effective amount" is an amount sufficient to achieve a clinically relevant change in the patient's condition through reduced expression of the targeted gene product,

most often a protein. In a preferred embodiment, the RNA molecule with intrinsic biological activity, *i.e.*, antisense RNA or ribozyme, will be expressed in transduced cells from its position-specific location in molar excess to the targeted RNA molecule. Expression levels of the heterologous and targeted RNAs can be determined by various assays, *e.g.*, by PCR analysis.

Typical dosages for *ex vivo* treatment of hematopoietic stem cells will generally range from about 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}, 10^{11} to 10^{12} infectious gene delivery vehicles, with dosages of 10^7 to 10^{10} infectious particles being preferred.

The gene delivery vehicles of the present invention may be administered to a wide variety of locations including, for example, into sites such as the cerebral spinal fluid, bone marrow, joints, arterial endothelial cells, rectum, buccal/sublingual, vagina, the lymph system, to an organ selected from the group consisting of lung, liver, spleen, skin, blood and brain, or to a site selected from the group consisting of tumors and interstitial spaces. Within other embodiments, the recombinant retrovirus may be administered intraocularly, intranasally, sublingually, orally, topically, intravesically, intrathecally, topically, intravenously, intraperitoneally, intracranially, intramuscularly, or subcutaneously. Other representative routes of administration include gastroscopy, ECRP and colonoscopy, which do not require full operating procedures and hospitalization, but may require the presence of medical personnel.

Considerations for administering the compositions of the present invention include the following:

Oral administration is easy and convenient, economical (no sterility required), safe (over dosage can be treated in most cases), and permits controlled release of the active ingredient of the composition (the gene delivery vehicle). Conversely, there may be local irritation such as nausea, vomiting or diarrhea, erratic absorption for poorly soluble drugs, and the gene delivery vehicle will be subject to "first pass effect" by hepatic metabolism and gastric acid and enzymatic degradation. Further, there can be slow onset of action, efficient plasma levels may not be reached, a patient's cooperation is required, and food can affect absorption. Preferred embodiments of the present invention include the oral administration of gene delivery vehicles that express genes encoding erythropoietin, insulin, GM-CSF cytokines, various polypeptides or peptide hormones, their agonists or antagonists, where these hormones can be derived from tissues such as the pituitary, hypothalamus, kidney, endothelial cells, liver, pancreas, bone, hemopoetic marrow, and adrenal. Such polypeptides can be used for induction of growth, regression of tissue, suppression of immune responses, apoptosis, gene expression, blocking receptor-ligand interaction, immune responses and can be treatment for certain anemias, diabetes, infections, high blood pressure, abnormal blood chemistry or chemistries (*e.g.*,

elevated blood cholesterol, deficiency of blood clotting factors, elevated LDL with lowered HDL), levels of Alzheimer associated amaloid protein, bone erosion/calcium deposition, and controlling levels of various metabolites such as steroid hormones, purines, and pyrimidines. Preferably, the gene delivery vehicles are first lyophilized, then filled into capsules and administered.

Buccal/sublingual administration is a convenient method of administration that provides rapid onset of action of the active component(s) of the composition, and avoids first pass metabolism. Thus, there is no gastric acid or enzymatic degradation, and the absorption of gene delivery vehicle is feasible. There is high bioavailability, and virtually immediate cessation of treatment is possible. Conversely, such administration is limited to relatively low dosages (typically about 10-15 mg), and there can be no simultaneous eating, drinking or swallowing. Preferred embodiments of the present invention include the buccal/sublingual administration of gene delivery vehicles that contain genes encoding self and/or foreign MHC, or immune modulators, for the treatment of oral cancer; the treatment of Sjogren's syndrome via the buccal/sublingual administration of such gene delivery vehicles that contain IgA or IgE antisense genes; and, the treatment of gingivitis and periodontitis via the buccal/sublingual administration of IgG or cytokine antisense genes.

Rectal administration provides a negligible first pass metabolism effect (there is a good blood/lymph vessel supply, and absorbed materials drain directly into the inferior vena cava), and the method is suitable of children, patients with emesis, and the unconscious. The method avoids gastric acid and enzymatic degradation, and the ionization of a composition will not change because the rectal fluid has no buffer capacity (pH 6.8; charged compositions absorb best). Conversely, there may be slow, poor or erratic absorption, irritation, degradation by bacterial flora, and there is a small absorption surface (about $0.05m^2$). Further, lipidophilic and water soluble compounds are preferred for absorption by the rectal mucosa, and absorption enhancers (*e.g.*, salts, EDTA, NSAID) may be necessary. Preferred embodiments of the present invention include the rectal administration of gene delivery vehicles that contain genes encoding colon cancer antigens, self and/or foreign MHC, or immune modulators.

Nasal administration avoids first pass metabolism, and gastric acid and enzymatic degradation, and is convenient. In a preferred embodiment, nasal administration is useful for gene delivery vehicle administration wherein the gene delivery vehicle carries nucleic acid capable of expressing a polypeptide with properties as described herein. Conversely, such administration can cause local irritation, and absorption can be dependent upon the state of the nasal mucosa.

Pulmonary administration also avoids first pass metabolism, and gastric acid and enzymatic degradation, and is convenient. Further, pulmonary administration permits localized actions that minimize systemic side effects and the dosage required for effectiveness, and there can be rapid onset of action and self-medication. Conversely, at times only a small portion of the administered composition reaches the brochioli/alveoli, there can be local irritation, and overdosing is possible. Further, patient cooperation and understanding is preferred, and the propellant for dosing may have toxic effects. Preferred embodiments of the present invention include the pulmonary administration of vehicles that express genes encoding IgA or IgE for the treatment of conditions such as asthma, hay fever, allergic alveolitis or fibrosing alveolitis, the CFTR gene for the treatment of cystic fibrosis, and protease and collagenous inhibitors such as a-1-antitrypsin for the treatment of emphysema. Alternatively, many of the same types of polypeptides or peptides listed above for oral administration may be used..

Ophthalmic administration provides local action, and permit prolonged action where the administration is via inserts. Further, avoids first pass metabolism, and gastric acid and enzymatic degradation, and permits self-administration via the use of eye-drops or contact lens-like inserts. Conversely, the administration is not always efficient, because the administration induces tearing. Preferred embodiments of the present invention include the ophthalmic administration of gene delivery vehicles that express genes encoding IgA or IgE for the treatment of hay fever conjunctivitis or vernal and atomic conjunctivitis; and ophthalmic administration of gene delivery vehicles that contain genes encoding melanoma specific antigens (such as high molecular weight-melanoma associated antigen), self and/or foreign MHC, or immune modulators.

Transdermal administration permits rapid cessation of treatment and prolonged action leading to good compliance. Further, local treatment is possible, and avoids first pass metabolism, and gastric acid and enzymatic degradation. Conversely, such administration may cause local irritation, is particularly susceptible to tolerance development, and is typically not preferred for highly potent compositions. Preferred embodiments of the present invention include transdermal administration to express genes encoding IgA or IgE for the treatment of conditions such as atopic dermatitis and other skin allergies; and transdermal administration of gene delivery vehicles encoding genes encoding melanoma specific antigens (such as high molecular weight-melanoma associated antigen), self and/or foreign MHC, or immune modulators.

Vaginal administration provides local treatment and one preferred route for hormonal administration. Further, such administration avoids first pass metabolism, and gastric acid and enzymatic degradation, and is preferred for administration of compositions wherein the gene delivery vehicles express peptides. Preferred

embodiments of the present invention include vaginal administration to express genes encoding self and/or foreign MHC, or immune modulators. Other preferred embodiments include the vaginal administration of genes encoding the components of sperm such as histone, flagellin, etc., to promote the production of sperm-specific antibodies and thereby prevent pregnancy. This effect may be reversed, and/or pregnancy in some women may be enhanced, by delivering gene delivery vehicles carrying vectors encoding immunoglobulin antisense genes, which genes interfere with the production of sperm-specific antibodies.

Intravesical administration permits local treatment for urogenital problems, avoiding systemic side effects and avoiding first pass metabolism, and gastric acid and enzymatic degradation. Conversely, the method requires urethral catheterization and requires a highly skilled staff. Preferred embodiments of the present invention include intravesical administration for delivering antitumor genes such as a prodrug activation gene such thymidine kinase or various immunomodulatory molecules such as cytokines.

Endoscopic retrograde cystopancreatography (ERCP) (goes through the mouth; does not require piercing of the skin) takes advantage of extended gastroscopy, and permits selective access to the biliary tract and the pancreatic duct. Conversely, the method requires a highly skilled staff, and is unpleasant for the patient.

Many of the routes of administration described herein (*e.g.*, into the CSF, into bone marrow, into joints, intravenous, intra-arterial, intracranial intramuscular, subcutaneous, into various organs, intra-tumor, into the interstitial spaces, intra-peritoneal, intralymphatic, or into a capillary bed) may be accomplished simply by direct administration using a needle, catheter or related device. In particular, within certain embodiments of the invention, one or more dosages may be administered directly in the indicated manner: into the cerebral spinal fluid at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; into bone marrow at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; into joint(s) at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; intravenously at dosages greater than or equal to 10^8, 10^9, 10^{10} or 10^{11} cfu; intra-arterially at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; intra-cranially at dosages greater than or equal to 10^9, 10^{10} or 10^{11} cfu; intra-muscularly at dosages greater than or equal to 10^{10} or 10^{11} cfu; intra-ocularly at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; pulmonarily at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; nasally at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; sub-lingually at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; rectally at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or

10^{11} cfu; orally at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; topically at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; vaginally at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; sub-cutaneously at dosages greater than or equal to 10^9, 10^{10} or 10^{11} cfu; inter-vesically at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; into an organ such as the lung, liver, spleen, skin, blood or brain at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; intra-tumor at dosages greater than or equal to 10^8, 10^9, 10^{10} or 10^{11} cfu; intra-peritoneally at dosages greater than or equal to 10^8, 10^9, 10^{10} or 10^{11} cfu; into interstitial spaces at dosages greater than or equal to 10^{10} or 10^{11} cfu; intra-lymphatically at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; into a capillary bed at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu; or intrathecally at dosages greater than or equal to 10^5, 10^6, 10^7, 10^8, 10^9, 10^{10} or 10^{11} cfu. "cfu" is understood to mean "colony forming unit" and is thus applicable to any of the various gene delivery vehicles useful in the practice of the invention. Colony forming unit means the number of cells transduced in vitro in some detectable fashion (*e.g.* drug resistance, gene expression detected by reaction with an antibody, PCR for transduced genes etc.)

Gene delivery vehicles may be delivered from outside of the organism to be treated, such as during a surgical procedure for other purposes, as part of a procedure with other purposes, or as a procedure designed expressly to administer the gene delivery vehicle. Other routes and methods for administration include the non-parenteral routes disclosed within U.S. Serial No. 08/366,788, filed December 30, 1994, as well as administration via multiple sites as disclosed within U.S. Serial No. 08/3__,___, filed December 30, 1994 [Atty. Docket No. 930049.427].

In addition to *in vivo* administration, the gene delivery vehicles of the invention may also be delivered in an *ex vivo* format.

The following examples are offered by way of illustration, and not by way of limitation.

Examples

The following examples are included to more fully illustrate the present invention. Additionally, these examples provide preferred embodiments of the invention and are not meant to limit the scope thereof. Standard methods for many of the procedures described in the following examples, or suitable alternative procedures, are provided in widely recognized manuals of molecular biology, such as, for example "Molecular Cloning," Second Edition (Sambrook, *et al.*, Cold Spring Harbor Laboratory Press, 1987) and "Current Protocols in Molecular Biology" (Ausubel, *et al.*, eds. Greene Associates/Wiley Interscience, NY, 1990).

EXAMPLE 1

PREPARATION OF A CHIMERIC RETROVIRAL INTEGRASE

This example describes how to confer the position specificity of the yeast Ty3 element (U.S. Patent No. 5,292,662) to the integrase (IN) protein derived from Moloney murine leukemia virus. Under appropriate conditions, the resulting chimeric integrase can then be packaged into a variety of gene delivery vehicles to confer on a vector construct according to the invention position-specific integration into the genome of a eukaryotic cell transduced or otherwise transformed therewith.

As noted previously, Ty3 IN is a functional and structural analog of retroviral IN proteins, including MoMLV IN. An alignment of the amino acid sequences of Ty3 and MoMLV (*see* FIG. 3) using the UWGCG Bestfit algorithm (Devereux, et al. (1984), Nucl. Acids Res., vol. 12:387-395) revealed roughly 25% identity within the core region which contains the D-D-E motif conserved among retrovirus IN proteins. In MoMLV IN, which comprises 408 amino acids, the core region corresponds to approximately 220 residues, the amino-terminal domain comprises about the first 40 amino acids, and the carboxy terminus the remaining 140 or so amino acids. In Ty3 IN, which is comprised of 536 amino acids, the amino-terminal domain spans about the first 60 residues, the core domain about the next 240 amino acids, with the carboxy-terminal domain comprising approximately 230 residues. In each protein, the amino-terminal domain is referred to as the "A" domain, the carboxyl end as the "C" domain, and the core domain as "B", as designated in FIG. 2. To further specify the origin of each domain, the single upper case designation therefor may be immediately followed by "m" for a domain derived from MoMLV IN or "t" for a domain derived from Ty3. For example, "At"

designates the A domain derived from Ty3, and a chimeric IN protein or gene therefor (or vector harboring such gene) referred to as AmBmCt, for example, is a chimera comprising the A and B domains from MoMLV IN and the C domain from Ty3. For purposes of designing each of the chimeric IN proteins described below and their corresponding genes, the actual A-B breakpoint, at residue 41 in MoMLV IN and 61 in Ty3, and B-C breakpoint, at amino acids 263 and 304 in MoMLV IN and Ty3, respectively, are selected so as to be within regions of random coil in order to minimize disruption of secondary and tertiary structure in the chimeric proteins.

A. Construction of Recombinant Retroviral Vectors Encoding Chimeric IN Proteins.

Each chimeric IN retroviral vector is based on the retroviral vector pRgpKan, which is illustrated in FIG. 4. pRgpKan contains all *cis* elements of a high titer retroviral vector. In addition, it expresses functional *gag* and *pol* gene products and, when introduced into a cell line expressing appropriate *env* gene products, results in the production of recombinant retroviral particles containing pRgpKan genomic RNA. pRgpKan contains the neomycin phosphotransferase gene from transposon Tn5; the gene is expressed as kanamycin resistance in bacteria and G418 resistance in mammalian cells (utilizing the SV40 early promoter). The retroviral vector also contains the colE1 origin of replication to allow the plasmid to be propagated in *E. coli.* The properties of pRgpKan allow for selection of transduced cells and subsequent rescue of vector DNA from those cells in bacteria after kanamycin selection. Genes expressed from pRgpKan-based retroviral vectors do not require the splice acceptor sequence found in the MoMLV IN gene since *env* is supplied *in trans*. The plasmid form of pRgpKan contains only one LTR. When expressed in cells, however, RNA genomes are transcribed with two LTRs. pRgpKan derived from the BAG vector (Price, et al. (1987), Proc.Natl.Acad.Sci.USA, vol 84:156-160.).

The seven chimeric retroviral vectors depicted in FIG. 5 were constructed using the pRgpKan backbone as follows. Initially, a 2.8 kb sequence containing the IN-coding region was excised from a plasmid such as pMLV-K (Miller *et al. J. Mol. Cell Biol. 5*:431, 1985) or 2xMLV which encodes the MoMLV genome, flanked on either end by a LTR (see FIG. 6). The excision was performed using *Sal* 1 and *Bam* H1. The resultant agarose gel purified fragment was cloned into *Sal* 1-*Bam* H1 digested pIBI-20 to generate pMLVIN (FIG. 6). Next, a 4.5kb *Sal* I to *Sca* I fragment from Ty3 containing the Ty3 IN-coding region

from pVB193, spanning base pairs 3132 - 5332 of the Ty3 element itself (Hansen, et al. (1990), J. Virol., vol. 64:2599-2607) was cloned into the gel purified 4.4 kb Sca I fragment of pMLVIN to yield the 8.8 kb plasmid vector pMLV/Ty3IN, which contained the heterologous MLV and Ty3 integrase genes, arranged in tandem as follows: AmBmCm-AtBtCt. Kunkel mutagenesis (Kunkel, T.A. (1985), Proc.Natl.Acad.Sci USA, vol. 82:488-492; looping out of single strand phage DNA using an oligonucleotide bridge complementary to the desired junction regions in MoMLV and Ty3 IN) was then performed on pMLV/Ty3IN to delete all but the first three codons of MLV IN in frame and immediately 5' of the Ty3 integrase to produce the plasmid pfTY3IN (FIG. 6). The nucleotide sequence of the oligonucleotide used to perform this mutagenesis, Oligo1 [SEQ ID NO: 5], is provided in FIG. 7. Also shown are the MLV and Ty3 sequences complementary to those of Oligo1.

pfTY3/MLVIN was then generated by ligating the 5.4 kb *Xmn* I to *Eco* RI fragment from pfTY3IN to the 3.8 Kb *Xmn* I fragment from pMLVIN. pfTY3/MLVIN contains the Ty3 and MLV integrases aligned AtBtCt-AmBmCm, as depicted in FIG. 6. As designed, the pfTY3/MLVIN construct contains the proteolytic cleavage site of MLV at the amino terminus of the Ty3-IN coding region. Each of the two phagemid constructs, pMLV/Ty3IN and pfTY3IN/MLVIN, were then used as substrates for Kunkel mutagenesis to generate four chimeric IN genes. Specifically, Oligo4 and Oligo5 [FIG. 7; SEQ ID NOS: 14 and 17, respectively] were used to generate pAmBtCt and pAmBmCt, respectively, from pMLV/Ty3IN. Oligo2 [FIG. 7; SEQ ID NO: 8] was used to mutagenize pfTY3/MLVIN to generate pAtBmCm, while Oligo3 [FIG. 7; SEQ ID NO: 11] was used to mutagenize pyTy3/MLVIN to generate pAtBtCm.

The last two chimeras, containing the B domain from one integrase between the A and C domains of the other were made as follows. pAtBtCm was cleaved with *Hind* III and *Msc* I to release the BtCm-containing region. This fragment was cloned into pAmBtCt which had been similarly cleaved to remove the BtCt region. The resultant construct, designated pAmBtCm, encoded a chimeric IN having A and C regions from MLV and a B domain from Ty3. pAtBmCt, which contains the A and C domain coding regions from Ty3 and the B domain coding region from MLV, was made by digesting pAmBmCt with *Hind* III and gel purifying the BmCt-containing region which was then cloned into pAtBmCm from which the BmCm region had been removed by *Hind* III digestion.

Each of the seven plasmids, or "phagemids," (pAmBmCt, pAtBmCm, pAtBmCt, pAmBtCm, pAtBtCm, pAmBtCt, and pfTy3IN) encoding the

chimeric IN proteins was then digested with either *Sal* 1-*Xho* 1 or *Sal* 1-*Bam* H1 to release the coding regions encoding the chimeric integrases. Each of the approximately 2.7 kb fragments containing the coding regions were then combined in separate three way ligations with the 4.1 kb *Nhe* I to *Sal* I fragment from pRgpKan and either a *Nhe* I to *Sal* I or *Nhe* I to *Bam* HI fragment of about 3 kb BAGÆX. *See* FIG. 4. The resultant constructs were designated pRgpAxByCz (where x, y, and z are either t or m, depending upon the phagemid source) depending on the origin of the A, B, and C domains in the chimeric integrase gene. Two independent kanamycin resistant clones for each construct were analyzed by restriction digest and confirmed by sequence analysis.

In addition to the seven chimeric retroviral vector constructs described above, a negative control retroviral vector, designated pRgpAmBtCt(-), was prepared by inserting the AmBtCt DNA sequence in the antisense orientation. When pRgpAmBtCt(-) is introduced into a suitable packaging cell line, *e.g.*, NC10 or 292 2-3, recombinant retroviral particles are produced, although these particles are replication incompetent and do not contain functional IN. pRgpKan, from which the MoMLV IN coding region was initially derived, was used as the positive control.

B. Production of Recombinant Retroviral Particles Using Recombinant Retroviral Vectors Coding for Chimeric IN Protein.

In order to test for biological activity and position-specific integration in human and other mammalian cells, cultures producing recombinant retroviral particles carrying chimeric integrase may be produced as described hereinafter. Initially, 293 2-3 cells (AKA 293-GP, Burns, *et al.*, *Proc.Natl.Acad.Sci. 90*:8033 1993, *see* WO 92/05266) generated from the human adenovirus 5-transformed embryonal kidney cell line 293 (ATCC #CRL1573) and which express MoMLV *gag* and *pol* genes, are produced. NC10 cells, which express the 4070A amphotropic envelope (*see* WO 92/05266), are generated from the HT1080 human fibrosarcoma line (ATCC #CCL 121). Both HT1080 and 292 cells have been demonstrated to lack DNA sequences that hybridize to the MoMLV genome. A plasmid construct, pMLP-G, expresses vesicular stomatitis virus G-protein and is used to complement MoMLV gag-pol proteins to pseudotype retroviral vector particles. FIG. 8 illustrates the method for generating producer cultures. In each case, the plasmid encoding the chimeric IN retroviral vector is co-transfected (Graham, et al. (1973), Virology, vol. 52:456-467) into 293 2-3 cells

along with pMLP-G at a 1:1 ratio. In the 293 2-3 cells, the retroviral RNA genome encoding the chimeric IN is encapsidated along with pol proteins into particles formed by *gag* gene products. These particles also contain VSV-G protein on their outer surface. After 48 hrs., filtered (.45 μm, cellulose acetate) supernatant fluid is placed onto NC10 cells. After an additional 24 hrs., the transduced (Emi, et al. (1991), J. Virol., vol. 65:1202-1207.) NC10 culture is placed under G418 selection (600 μg/ml). Selection continues until non-transduced NC10 control cultures no longer contain viable cells. The resulting cultures produce retroviral particles which consist of: a RNA genome coding for, among other things, a chimeric IN protein; *gag* and *pol* gene products expressed from the chimeric retroviral vector construct; an amphotropic envelope; and chimeric IN protein.

Functionality of these retroviral particles is subsequently assessed by placing filtered supernatant fluid onto target cells, *e.g.*, HT1080, and selecting for G418 resistant transductants. All of the chimeric retroviral vector particles produced have been used successfully to generate NC10 producer cells, *i.e.*, cells which produce infectious recombinant retroviral particles. As judged by relative transient titers, retroviral genomes coding for a chimeric IN protein are packaged at the same rate as observed for RgpKan.

To further characterize the NC10 producer cells to ensure that the biological activity of the *gag* and *pol* components, NC10 cell lysates and supernatants are tested by Northern analysis (Sambrook, et al., *supra*) to determine the level of expression and packaging of the retroviral constructs coding for the chimeric integrases, as compared to the expression levels detected for the control RgpKan construct. Because no *cis* acting sequences required for transcription, packaging, or replication are modified, the results for the chimeric retroviral vectors are expected to be similar to those determined for the control.

Additionally, NC10 cell lysates and supernatants are tested by Western blot analysis for production of MoMLV *gag* and IN proteins using anti-p30, anti-MoMLV IN, and anti-Ty3 IN rabbit polyclonal antibodies prepared separately according to standard techniques. Because the levels of gag protein is comparable for control RgpKan and the seven chimeric constructs, they serve as an internal control for the levels of *pol*-derived IN protein. The mobility and detected amount of chimeric integrase produced from each construct indicates if correct processing by the MoMLV protease occurred at the MoMLV IN processing site three amino acids in front of the junctions with the Ty3 sequence, for those chimeras containing an At domain. As retroviral processing sites are fairly locally determined, requiring approximately seven residues (Pettit, et al. (1991), J. Biol. Chem., vol.

266:14539-14547) and are similar to Ty3 protease processing sites, functional processing of the MoMLV IN site should occur even for the AtBtCt construct.

In addition, NC10 cell supernatants are concentrated by centrifugation and tested for reverse transcriptase activity (Goff, et al. (1981), J. Virol., vol. 38:329-248.) to determine if incorporation of the heterologous IN domain into the retroviral particle disrupts reverse transcriptase function. Activity of test constructs is compared to activity in the RgpKan supernatants and both are normalized to levels of gag protein.

For each of the two clones of the seven chimeric IN constructs and two control constructs, RgpKan and RgpAmBtCt(-), three G418 resistant cultures are derived and supernatants are harvested.

C. Target Cell Transduction.

In order to evaluate whether the recombinant retroviral vectors generated from NC10 producers, as described above, are transduction competent, a transduction assay using a eukaryotic cell line, *e.g.*, HT1080, is employed. HT1080 cells are susceptible to infection by amphotropic retroviral vectors, however, here transduction (as determined by G418 resistance) depends on successful integration mediated by a chimeric IN incorporated into the retroviral particle during virion assembly in the NC10 producer cell, where the integrase function is encoded in NC10 cells by the retroviral vector introduced by the VSV-g pseudotype virion produced in the 292 2-3 cells.

Transduction assays are performed as follows: Approximately 10 ml of filtered (0.45 μm) supernatant fluid from NC10 producers is placed on about 10^6 fresh HT1080 cells. After 24 hrs., the media is replaced with DMEM + 10% FBS. After an additional 24 hrs., the media is changed again, and this time contains 600 μg/ml G418. When non-transduced control cultures of HT1080 cells are no longer viable under G418 selection, G418 resistant colonies are scored as successful transduction events. The number of HT1080 G418-resistant clones derived from the chimeric constructs, as well as from the control constructs, are then normalized on the basis of the amount of *gag* protein in the NC10 supernatant, as measured by Western blot, so that the relative infectivity of control and chimeric constructs is normalized for particle production. The results of a transduction assay using the chimeric constructs described above appear in Table 1, below, and indicate that the substitution of Ty3 sequences for the corresponding domains of MoMLV IN result in an integrase capable of mediating integration.

Table 1

Transduction Assay Results

CONSTRUCT	ISOLATE#	Titer on HT1080 A	B
RgpAmBmCt	4-11	2	³3
RgpAtBmCm	29-11		³6
RgpAtBmCt	3-49	0	0
RgpAmBtCm	2-6	>500	³500
RgpAtBtCm	1-8	>400	³100
RgpAmBtCt	24-24	2	0
RgpfTy3		0	³1
RgpKan		$>10^4$	$>10^4$
pRgpAmBtCt(-)			0

To analyze reverse transcription and viral DNA processing, four hrs. after infection, cells are harvested and Hirt supernatants derived (Roth, et al. (1989), Cell, vol. 58:47-54.). The amount of supernatant used for each sample is normalized on the basis of gag protein levels. These supernatants should contain extrachromosomal viral DNA produced by reverse transcription of the retroviral RNA genome following cell entry. After RNase digestion, the level of reverse transcribed DNA as determined by Southern blot analysis using MoMLV-specific probes. Viral DNA processing is assayed by examination of the ends of the extrachromosomal DNA. Both retrovirus and Ty3 IN proteins remove two bp from the 3' ends of replicated retroviral DNA prior to integration. DNA in the Hirt supernatants fluid is digested with restriction enzymes close to the ends of the virus, fractionated on a gel suitable for resolving fragments of the expected size (which is dependent on the sequence of the nucleic acid and the restriction enzyme selected), transferred to nitrocellulose and probed with terminal strand specific probes to determine whether terminal processing has occurred. This information will be used as described below in (Section D) to optimize another set of constructs.

D. Analysis of Position-Specific Integration

To determine the position of provirus in a eukaryotic genome, an integration library may be constructed from DNA of transduced cells. Below, such

an analysis is described for provirus whose integration into the genomes of HT1080 target cells was mediated by the chimeric IN proteins described above. Each of the retroviral vectors encoding the chimeric integrases and the pRgpKan control contain the col E1 origin of replication (ori), as well as bacterial control elements that allow the expression of the kanamycin resistance gene in *E. coli*. As a result, the integrated provirus, along with host DNA adjacent to the site of integration, can be cloned directly in bacteria (Cepko, et al. (1984), Cell, vol. 37:1053-1062).

FIG. 9 illustrates the technique described herein for isolating integrated proviral DNA. After HT1080 target cells are selected for transduction by G418 treatment, genomic DNA from pooled cultures is isolated and digested to completion with restriction enzymes that do not occur in the proviral DNA, *e.g.*, *Sca* 1, *Bst* 1107I, for which no restriction recognition site occurs in the chimeric vector constructs. The DNA is then ligated under conditions that favor circularization so that host DNA on either side of the integration site is included in the circular molecules. Circular DNA is then used to transform the DH12s C strain of *E. coli* (Gibco BRL), or another strain of *E. coli* designed for cloning non-bacterial DNA. Kanamycin resistant colonies are then isolated and pooled, followed by plasmid DNA extraction.

As Ty3 integrase mediates integration proximal to RNA polymerase III transcribed genes, which include tRNA genes, to determine the approximate frequency of insertions next to tRNA genes, animal cell tRNA is purified, labeled (Goddard, et al. (1983), Nuc. Acids Res., vol. 11:2551-62.), and used to probe Southern blots of pooled plasmid DNA from the kanamycin resistant bacteria. Parallel blots probed with LTR specific probes reveal the relative abundance of proviral DNA and adjacent tRNA genes. Nonspecific integration, as expected for RgpKan controls, yields a frequency of association provirus and a tRNA gene less than 1 in 1,000 clones, based on an average size of cloned integration fragment of 3 kb and a total genomic DNA of 10^6 kb harboring about 1,300 tRNA genes (Hatlen, et al. (1971), J. Mol. Biol., vol. 56:535-553). Pooled DNA and colonies can also be screened in a similar manner using 5S and U6 sequence probes, since these genes are conserved in sequence but repeated in the genome, 1,000 to 2,000 for the 5S gene (Sorensen, et al. (1991), Nuc. Acids Res., vol. 19, 4147-51) and 200 for U6 (Hayashi, K. (1981), Nuc. Acids Res., vol. 9:3379-88).

Clones and pools from the integration library may also be probed with a radiolabeled Alu-specific probe. Although Alu elements are generally not transcriptionally active, they contain pol III promoter elements.

Preferably, at least ten plasmids for each construct are analyzed by sequencing to identify potential target genes independent of probe reactivity (Pavesi, et al. (1994), Nuc. Acids Res., vol. 22:1247-56). Sequencing from the LTRs allows a determination of provirus orientation with respect to the pol III-transcribed element. If an insertion is within 500 bp, preferably within 100 bp, more preferably within 20 bp of the 5' end of the mature tRNA coding region (or coding region for another RNA pol III gene), integration is deemed to have involved position-specificity. However, only after approximately 10-20 position-specific events are analyzed is a chimeric IN function be considered position-specific.

In addition, the level of expression of a gene of interest from a provirus integrated in a position-specific manner by a chimeric integrase according to the invention can also be analyzed. Preferably, expression levels will vary less than 100-fold, more preferably less than 50-fold, and most preferably less than 10-fold for a given gene of interest integrated in a position-specific manner, as compared to the same gene randomly integrated in a comparable eukaryotic cell genome. Typically, target cells are tested by quantitative Northern blot analysis for the expression levels of transcripts produced from the integrated provirus. To perform a statistical analysis, at least ten independent clones are examined per original construct.

EXAMPLE 2

PREPARATION OF RETROVIRAL VECTOR BACKBONES

A. Preparation of Retroviral Backbones KT-1 and KT-3B

The Moloney murine leukemia virus (MoMLV) 5' long terminal repeat (LTR) EcoR I-EcoR I fragment, including *gag* sequences, from the N2 vector (Armentano *et al.*, *J. Vir. 61*:1647-1650, 1987; Eglitas *et al.*, *Science 230*:1395-1398, 1985) is ligated into the plasmid SK^+ (Stratagene, La Jolla, CA). The resulting construct is designated N2R5. The N2R5 construct is mutated by site-directed *in vitro* mutagenesis to change the ATG start codon to ATT preventing gag expression. This mutagenized fragment is 200 base pairs (bp) in length and flanked by Pst I restriction sites. The Pst I-Pst I mutated fragment is purified from the SK^+ plasmid and inserted into the Pst I site of N2 MoMLV 5' LTR in plasmid pUC31 to replace the non-mutated 200 bp fragment. The plasmid pUC31 is derived from pUC19 (Stratagene, La Jolla, CA) in which additional restriction sites Xho I, Bgl II, BssH II and Nco I are inserted between the EcoR I and Sac I sites of the polylinker. This construct is designated pUC31/N2R5gM.

A 1.0 kilobase (Kb) MoMLV 3' LTR EcoR I-EcoR I fragment from N2 is cloned into plasmid SK^+ resulting in a construct designated N2R3$^-$. A 1.0 Kb Cla I-Hind III fragment is purified from this construct.

The Cla I-Cla I dominant selectable marker gene fragment from pAFVXM retroviral vector (Kriegler *et al., Cell 38*:483, 1984; St. Louis *et al., PNAS 85*:3150-3154,1988), comprising a SV40 early promoter driving expression of the neomycin (neo) phosphotransferase gene, is cloned into the SK^+ plasmid. This construct is designated SK^+ SV_2-*neo* A 1.3 Kb Cla I-BstB I gene fragment is purified from the SK^+ SV_2-*neo* plasmid.

KT-3B or KT-1 vectors are constructed by a three part ligation in which the Xho I-Cla I fragment containing the gene of interest and the 1.0 Kb MoMLV 3' LTR Cla I-Hind III fragment are inserted into the Xho I-Hind III site of pUC31/N2R5gM plasmid. This gives a vector designated as having the KT-1 backbone. The 1.3 Kb Cla I-BstB I *neo* gene fragment from the pAFVXM retroviral vector is then inserted into the Cla I site of this plasmid in the sense orientation to yield a vector designated as having the KT-3B backbone.

EXAMPLE 3

CONSTRUCTION OF RECOMBINANT RETROVIRUSES EXPRESSING FACTOR VIII

A. Construction of Full-Length and B Domain Deleted Factor VIII cDNA Retroviral Vector

The following is a description of the construction of several retroviral vectors encoding a full-length factor VIII cDNA. Further discussion is also provided in U.S. Serial No. 08/366,851, filed December 30, 1994. Due to the packaging constraints of retroviral vectors and because selection for transduced cells is not a requirement for therapy, a retroviral backbone, *e.g.*, KT-1, lacking a selectable marker gene is employed.

1. Production of Plasmid Vectors Encoding Full-Length Factor VIII

A gene encoding full length factor VIII can be obtained from a variety of sources. One such source is the plasmid pCIS-F8 (*see* EP 0 260 148), which contains a full length factor VIII cDNA whose expression is under the control of a CMV major immediate-early (CMV MIE) promoter and enhancer. The factor VIII cDNA contains approximately 80 bp of 5' untranslated sequence from the factor VIII gene and a 3' untranslated region of about 500 bp. In addition, between the CMV promoter and the

factor VIII sequence lies a CMV intron sequence, or "cis" element. The cis element, spanning about 280 bp, comprises a splice donor site from the CMV major immediate-early promoter about 140 bp upstream of a splice acceptor from an immunoglobulin gene.

More specifically, a plasmid, designated pJW-2, encoding a retroviral vector for expressing full length factor VIII is constructed using the KT-1 backbone from pKT-1. Briefly, in order to facilitate directional cloning of the factor VIII cDNA insert into pKT-1, the unique Xho I site is converted to a Not I site by site directed mutagenesis. The resultant plasmid vector is then opened with Not I and Cla I. pCIS-F8 is digested to completion with Cla I and Eag I, for which there are two sites, to release the fragment encoding full length factor VIII. This fragment is then ligated into the Not I/Cla I restricted vector to generate a plasmid designated pJW-2.

2. Construction of a Truncated Factor VII Retroviral Vector (ND-5)

A plasmid vector encoding a truncation of about 80% (approximately 370 bp) of the 3' untranslated region of the factor VIII cDNA, designated pND-5, is constructed in a pKT-1 vector as follows: As described for pJW-2, the pKT-1 vector employed has its Xho I restriction site replaced by that for Not I. The factor VIII insert is generated by digesting pCIS-F8 with Cla I and Xba I, the latter enzyme cutting 5' of the factor VIII stop codon. The approximately 7 kb fragment containing all but the 3' coding region of the factor VIII gene is then purified. pCIS-F8 is also digested with Xba I and Pst I to release a 121 bp fragment containing the gene's termination codon. This fragment is also purified and then ligated in a three way ligation with the larger fragment encoding the rest of the factor VIII gene and Cla I/Pst I restricted BLUESCRIPT® KS^+ plasmid (Stratagene, *supra*) to produce a plasmid designated pND-2.

The unique Sma I site in pND-2 is then changed to a Cla I site by ligating Cla I linkers (New England Biolabs, Beverly, MA) under dilute conditions to the blunt ends created by a Sma I digest. After recircularization and ligation, plasmids containing two Cla I sites are identified and designated pND-3.

The factor VIII sequence in pND-3, bounded by Cla I sites and containing the full length gene with a truncation of much of the 3' untranslated region, is cloned as follows into a plasmid backbone derived from a Not I/Cla I digest of pKT-1 (a pKT-1 derivative by cutting at the Xho I site, blunting with Klenow, and inserting a Not I linker (New England Biolabs)), which yields a 5.2 kb Not I/Cla I fragment. pCIS-F8 is cleaved with Eag I and Eco RV and the resulting fragment of about 4.2 kb, encoding the 5' portion of the full length factor VIII gene, is isolated. pND-3 is digested with Eco RV and Cla I and a 3.1 kb fragment is isolated. The two fragments containing portions of the factor VIII

gene are then ligated into the Not I/Cla I digested vector backbone to produce a plasmid designated pND-5.

3. Construction of the B-Domain Deleted Vector

The precursor DNA for the B-deleted FVIII is obtained from Miles Laboratory. This expression vector is designated p25D and has the exact backbone as pCISF8 above. The Hpa I site at the 3' of the FVIII8 cDNA in p25D is modified to Cla-I by oligolinkers. An Acc I to Cla I fragment is clipped out from the modified p25D plasmid. This fragment spans the B-domain deletion and includes the entire 3' two-thirds of the cDNA. An Acc I to Cla I fragment is removed from the retroviral vector JW-2 above, and replaced with the modified B-domain deleted fragment just described. This is designated B-del-1.

The vectors described herein are then used to produce infectious, replication incompetent recombinant retroviral particles incorporating a chimeric integrase protein in an appropriate retroviral packaging cell line, preferably a human packaging cell line.

Detection of Replication Competent Retroviruses (RCR)

1. The Extended S^+L^- Assay

The extended S^+L^- assay determines whether replication competent, infectious virus is present in the supernatant of the cell line of interest. The assay is based on the empirical observation that infectious retroviruses generate foci on the indicator cell line $MiCl_1$ (ATCC No. CCL 64.1). The $MiCl_1$ cell line is derived from the Mv1Lu mink cell line (ATCC No. CCL 64) by transduction with Murine Sarcoma Virus (MSV). It is a non-producer, non-transformed, revertant clone containing a replication defective murine sarcoma provirus, S^+, but not a replication competent murine leukemia provirus, L^-. Infection of $MiCl_1$ cells with replication competent retrovirus "activates" the MSV genome to trigger "transformation" which results in foci formation.

Supernatant is removed from the cell line to be tested for presence of replication competent retrovirus and passed through a 0.45 μ filter to remove any cells. On day 1, Mv1Lu cells are seeded at 1.0×10^5 cells per well (one well per sample to be tested) of a 6 well plate in 2 ml DMEM, 10% FBS and 8 μg/ml polybrene. Mv1Lu cells are plated in the same manner for positive and negative controls on separate 6 well plates. The cells are incubated overnight at 37°C, 10% CO_2. On day 2, 1.0 ml of test supernatant is added to the Mv1Lu cells. The negative control plates are incubated with

1.0 ml of media. The positive control consists of three dilutions (200 focus forming units (ffu), 20 ffu and 2 ffu each in 1.0 ml media) of MA virus (referred to as pAM in Miller *et al., Molec. and Cell Biol. 5*:431, 1985) which is added to the cells in the positive control wells. The cells are incubated overnight. On day 3, the media is aspirated and 3.0 ml of fresh DMEM and 10% FBS is added to the cells. The cells are allowed to grow to confluency and are split 1:10 on day 6 and day 10, amplifying any replication competent retrovirus. On day 13, the media on the Mv1Lu cells is aspirated and 2.0 ml DMEM and 10% FBS is added to the cells. In addition, the $MiCl_1$ cells are seeded at 1.0×10^5 cells per well in 2.0 ml DMEM, 10% FBS and 8 µg/ml polybrene. On day 14, the supernatant from the Mv1Lu cells is transferred to the corresponding well of the $MiCl_1$ cells and incubated overnight at 37°C, 10% CO_2. On day 15, the media is aspirated and 3.0 ml of fresh DMEM and 10% FBS is added to the cells. On day 21, the cells are examined for focus formation (appearing as clustered, refractile cells that overgrow the monolayer and remain attached) on the monolayer of cells. The test article is determined to be contaminated with replication competent retrovirus if foci appear on the $MiCl_1$ cells. Using these procedures, it can be shown that the HBV core producer cell lines are not contaminated with replication competent retroviruses.

2. Cocultivation of Producer Lines and MdH Marker Rescue Assay

As an alternate method to test for the presence of RCR in a vector-producing cell line, producer cells are cocultivated with an equivalent number of *Mus dunni* (NIH NIAID Bethesda, MD) cells. Small scale cocultivations are performed by mixing of 5.0×10^5 *Mus dunni* cells with 5.0×10^5 producer cells and seeding the mixture into 10 cm plates (10 ml standard culture media/plate, 4 µg/ml polybrene) at day 0. Every 3-4 days the cultures are split at a 1:10 ratio and 5.0×10^5 *Mus dunni* cells are added to each culture plate to effectively dilute out the producer cell line and provide maximum amplification of RCR. On day 14, culture supernatants are harvested, passed through a 0.45 µ cellulose-acetate filter, and tested in the MdH marker rescue assay. Large scale cocultivations are performed by seeding a mixture of 1.0×10^8 *Mus dunni* cells and 1.0×10^8 producer cells into a total of twenty T-150 flasks (30 ml standard culture media/flask, 4 µg/ml polybrene). Cultures are split at a ratio of 1:10 on days 3, 6, and 13 and at a ratio of 1:20 on day 9. On day 15, the final supernatants are harvested, filtered and a portion of each is tested in the MdH marker rescue assay.

The MdH marker rescue cell line is cloned from a pool of *Mus dunni* cells transduced with LHL, a retroviral vector encoding the hygromycin B resistance gene (Palmer *et al., PNAS 84*: 1055-1059, 1987). The retroviral vector can be rescued from MdH cells upon infection of the cells with RCR. One ml of test sample is added to a well

of a 6-well plate containing 10^5 MdH cells in 2 ml standard culture medium (DMEM with 10% FBS, 1% 200 mM L-glutamine, 1% non-essential amino acids) containing 4 µg/ml polybrene. Media is replaced after 24 hours with standard culture medium without polybrene. Two days later, the entire volume of MdH culture supernatant is passed through a 0.45 µ cellulose-acetate filter and transferred to a well of a 6-well plate containing 5.0×10^4 *Mus dunni* target cells in 2 ml standard culture medium containing polybrene. After 24 hours, supernatants are replaced with standard culture media containing 250 µg/ml of hygromycin B and subsequently replaced on days 2 and 5 with media containing 200 µg/ml of hygromycin B. Colonies resistant to hygromycin B appear and are visualized on day 9 post-selection, by staining with 0.2% Coomassie blue.

Administration of Vector Construct

1. Animal Administration Protocol

The intestinal epithelium is an attractive site for gene delivery due to its rapidly proliferating tissue mass and the known location of stem cells in the crypts of Lieberkuhn. The deep location of the stem cells in the crypts and the protective role of the mucus gel layer, makes the retrovirus relatively inaccessible to the tissue cells. However, the accessibility of the retroviral vector to these stem cells can be improved in animal models by the *in vivo* mucus removal method of Sandberg, J., *et al.*, (*Human Gene Therapy* 5:3232-329, 1994).

Male Sprague-Dawley rats obtained from Charles River Breeding Laboratories (Portage, MD.) are anesthetized and the cecum is identified upon opening the peritoneal cavity. A 3 cm ileal segment is isolated from the last Peyer's patch in the terminal ileum and ligated at each end. A plastic catheter attached to a syringe is inserted into the segment and two milliliters of the mucolytic agents dithiothreitol and N-acetyl-cysteine is instilled under mild pressure for two minutes, then removed. This procedure is repeated once again before filling the segment with 0.2 to 2.0 ml of retroviral vector particles at 10^6 to 10^{10} cfu/ml. The ligatures are removed 1 to 4 hours later and the abdominal cavity is sutured. Control animals are instilled with formulation buffer only.

Blood is collected from the tail vein and assayed for factor VIII production by a sandwich ELISA specific for human factor VIII (according to the modified procedure of Zatloukal, K., *et al.*, *PNAS 91*:5148-5152, 1994). The ELISA is based on two monoclonal antibodies directed against human factor VIII (ESH 4 and ESH 8: American Diagnostica). ESH 4 (25 mg/ml in 1.0 M $NaHCO_3$/0.5 M NaCl, pH 9.0) is coupled to the ELISA plates overnight at 4°C, washed with 0.1% Tween 20 in PBS, and blocked with 1% BSA in PBS. The samples are applied in 0.05 M Tris-HCl/1 M

NaCl/2% BSA, pH 7.5, over 4 hr at room temperature, the plates are washed, and ESH 8 (2.5 mg/ml in 0.05 M Tris-HCl/1 M NaCl/2% BSA, pH 7.5,) which has been biotinylated with *N*- hydroxysuccinimidobiotin (Pierce, Rockford, IL.) is added for 2 hr at room temperature. The color reaction is performed with peroxidase-conjugated streptavidin (Boehringer Mannheim, Indianapolis, IN.) and o- phenylenediamine dihydrochloride as substrate. The human factor VIII:c standard (from the National Institute for Biological Standards and Control, Hertfordshire, U.K.) and normal rat plasma are used as references.

2. Human Administration Protocol

Lyophilized recombinant retrovirus containing the gene for Factor VIII expression is formulated into an enteric coated tablet or gel capsule according to known methods in the art. These are described in the following patents: US 4,853,230, EP 225,189, AU 9,224,296, AU 9,230,801, and WO 92/14452.

The capsule is administered orally to be targeted to the jejunum. At 1 to 4 days following oral administration of the recombinant retrovirus, expression of Factor VIII is measured in the plasma and blood by the Coatest^{R} Factor VIII assay as described in Example 2B1.

EXAMPLE 4

INTRAVESICAL ADMINISTRATION OF RECOMBINANT RETROVIRUSES EXPRESSING TK

Construction of TK Vector Constructs

1. Construction of plasmids containing vector LTR sequences

All of the following retroviral vectors are based on the N2 vector (Keller et al., *Nature 318*:149-154, 1985). Briefly, 5' and 3' Eco RI LTR fragments (2.8 and 1.0 Kb, respectively) (Armentano, *J. Vir. 61*:1647, 1987; Eglitis, *Science 230*:1395, 1985) are initially subcloned into the Eco RI site of plasmids SK^{+} (Stratagene, San Diego, CA) and pUC31. pUC31 is a modification of pUC19 (Stratagene, San Diego, CA) carrying additional restriction sites (Xho I, Bgl II, BssH II, and Nco I) between the Eco RI and Sac I sites of the polylinker. Plasmid N2R3+/- is thereby created from ligation of the SK^{+} plasmid with the 1.0 Kb 3' LTR fragment. The plasmids p3 1N2R5+/- and p3 1N2R3+/- are constructed from the ligation of pUC31 with the 2.8 Kb 5' LTR and packaging signal (Y) or the 1.0 Kb 3' LTR fragment, respectively. In each case N2 refers to the vector source, R refers to the fact that the fragment is an Eco RI fragment, 5 and 3 refer to 5' or

3' LTRs, and + or - refers to the orientation of the insert (see Figures 1-6 for examples of LTR subclones).

In one case, a 1.2 Kb Cla I/Eco RI 5' LTR and W fragment from N2 is subcloned into the same sites of an SK^+ vector. This vector is designated pN2CR5. In another case, the 5' LTR containing a 6 bp deletion of the splice donor sequence (Yee *et al.*, Cold Spring Harbor, Quantitative Biology, 51:1021, 1986) is subcloned as a 1.8 Kb Eco RI fragment into pUC31. This vector is designated p31N25D[+], Figure 6.

2. Construction of plasmids containing HSVTK

The coding region and transcriptional termination signals of HSV-1 thymidine kinase gene (HSVTK) are isolated as a 1.8 Kb Bgl II/Pvu II fragment from plasmid 322TK (3.5 kb Bam HI fragment of HSV-1 (McKnight *et al.*) cloned into Bam HI of pBR322 (ATCC No. 31344)) and cloned into Bgl II/Sma I-digested pUC31. This construct is designated pUCTK. For constructs which require deletion of the terminator signals, pUCTK is digested with Sma I and Bam HI and the 0.3 Kb fragment containing the $(A)_n$ signal is removed. The remaining coding sequences and sticky-end Bam HI overhang are reconstituted with a double-stranded oligonucleotide made from the following oligomers:

(SEQUENCE ID. NO. 18)

5' GAG AGA TGG GGG AGG CTA ACT GAG 3'

(SEQUENCE ID. NO. 19)

5' GAT CCT CAG TTA GCC TCC CCC ATC TCT C 3'

The resulting construct is designated pTKD A, Figure 7.

For diagnostic purposes, the oligonucleotides are designed to destroy the Sma I site while maintaining the Ava I site without changing the translated protein.

The plasmid pPrTKDA (Figure 8), which contains the HSVTK promoter and coding sequence (lacking an $(A)_n$ signal), is constructed as follows.

1. pTKD A is linearized with Bgl II treated with alkaline phosphatase, and gel purified.

2. A 0.8 Kg fragment contained the HSVTK transcriptional promoter is isolated as a Bam HI/Bgl II fragment from p322TK.

3. Products from (1) and (2) are ligated, transformed into bacteria, and positive clones are screened for the proper orientation of the promoter region. A resultant clone is designated pPrTKDA (Figure 8).

3. Construction of retroviral provectors expressing HSVTK from a constitutive promoter

The retroviral provectors pTK-1 and pTK-3 are constructed essentially as described below.

1. The 5 Kb Xho I/Hind III 5' LTR and plasmid sequences are isolated from p31N2R5(+) (Figure 1).

2. HSVTK coding sequences lacking transcriptional termination sequences are isolated as a 1.2 Kb Xho I/Bam HI fragment from pTKDA (Figure 2).

3. 3' LTR sequences are isolated as a 1.0 Kb Bam HI/Hind III fragment from pN2R3(-) (Figure 2).

4. The fragments from steps 1-3 are mixed, ligated, transformed into bacteria, and individual clones identified by restriction enzyme analysis. The construct is designated TK-1 (Figure 9).

5. pTK-3 is constructed by linearizing TK-1 with Bam HI, filling in the 5' overhang and blunt-end ligating a 5'-filled Cla I/Cla I fragment containing the bacterial lac UV5 promoter, SV40 early promoter, plus Tn5 neo^r gene obtained from pAFVXM retroviral vector (Krieger *et al.*, *Cell 39*:483, 1984; St. Louis *et al.*, *PNAS 85*:3150, 1988). Kanamycin-resistant clones are isolated and individual clones are screened for the proper orientation by restriction enzyme analysis (Figure 9).

These constructs were used to generate infectious recombinant vector particles in conjunction with a packaging cell line, such as DA as described above.

EXAMPLE 5

PREPARATION OF RECOMBINANT RETROVIRUS FOR DELIVERY OF HUMAN GROWTH HORMONE

A. Preparation of hGH containing vectors

Vector pDHF828 containing the full-length human growth hormone gene is constructed essentially as follows. Briefly, plasmid pDHF811, was constructed by removing the XhoI- ClaI fragment of the KT-1 retroviral vector described above, and inserting the following oligonucleotide linkers by ligation of the cohesive ends:

Linker sequences:
(SEQUENCE ID# 20)

5' TCGAGGATCC GCCCGGGCGG CCGCATCGAT GTCGACG 3'

(SEQUENCE ID# 21)

5' CGCGTCGA CATCGATGCG GCCGCCCGGG CGGATCC 3'

In particular, the linkers were annealed at 65°C for 20 minutes, 42°C for 20 minutes, 37°C for 20 minutes, and room temperature for 2 hours. The concentrations of both oligonucleotides was 18mM and the salt concentration was 100 mM NaCl. After annealing, 50ml of 1.8 mM annealed linker was digested with ClaI overnight to generate ClaI ends. For ligation, 3nM of KT-1 XhoI - ClaI fragment was mixed with 90nM of linker, and the resultant mixture incubated at 15°C for 3 hours. The ligated DNA sample was transformed into DH-5a competent cells, followed by screening of transformants.

Plasmid chGH 800 containing the full length cDNA of the hGH gene (Martial, R.A. et al., *Science 205*:602, 1979) was digested with Hind III, blunt-ended with the Klenow fragment enzyme, and cloned into the SrfI site of pDHF811. The resultant plasmid was designated pDHF828 and can be introduced into an appropriate packaging cell line to produce recombinant retroviral particles incorporating a chimeric integrase protein which confers position-specific integration into the targeted genome.

EXAMPLE 6

ANALYSIS OF CRUDE AND PURIFIED RECOMBINANT RETROVIRUS

Crude and purified solutions of recombinant retrovirus particles may be separated on gradient polyacrylamide gels utilizing, for example, the PHASTGEL system (Pharmacia Biotech). Briefly, samples are place on 4-15% polyacrylamide gels without pretreatment and electrophoresed for 35 minutes at 250V. The gels are then removed and stained with coomassie blue in order to detect virus and other protein components. The gels are then scanned by laser densitometry in order to determine the content of virus and other components.

Virus bands may be identified by their relative molecular weight and by reverse transcriptase activity (RT). The purpose of this assay is to quantify the activity of reverse transcriptase (RT), an enzyme exclusively associated with all retroviruses. The relative amount of retrovirus in a sample can be determined by measuring the activity of this enzyme in a given preparation.

Briefly, Moloney murine leukemia virus reverse transcriptase (Pharmacia, Newark, NJ) is diluted to a concentration of 1 μg/ml by addition of 1x Tris/EDTA buffer solution containing 10 mM Tris-HCl and 1mM EDTA, pH 8.0. One hundred microliters

of this solution is added 6.84 ml of sterile dH_20, 500 µl of 1M Tris HCl pH 8.0, 10 µl of 0.1M $MnCl_2$, 200 µl of 1M dithiothreitol, 50 µl of 10% Nonidet P40 (NP40), 2 µl of 100 µM dNTP (Pharmacia, Newark, NJ, dNTP Ultrapure Kit™), and 300 µl Methyl - ^3H Thymidine 5' - Triphosphate (30-50 Ci/mmol). This mixture is incubated for 1 hour at 37°C in a water bath. Following incubation the sample is placed on ice. Approximately 1.0 ml of 2N HCl is added to the cooled sample. The precipitated radiolabled DNA fragments are vacuum filtered onto glass fiber filters using a Millipore sampling manifold (Millipore, Philadelphia, PA). The filters are washed, dried, placed in scintillation cocktail, and counted in a Beckman LS5000TD scintillation counter (Beckman, Dallas, TX).

* * *

While the present invention has been described above both generally and in terms of preferred embodiments, it is understood that variations and modifications will occur to those skilled in the art in light of the description, *supra*. Therefore, it is intended that the appended claims cover all such variations coming within the scope of the invention as claimed.

Additionally, the publications and other materials cited to illuminate the background of the invention, and in particular, to provide additional details concerning its practice as described in the detailed description and examples, are hereby incorporated by reference in their entirety.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: CHIRON VIAGENE, INC.

(ii) TITLE OF INVENTION: POSITION-SPECIFIC INTEGRATION OF VECTOR CONSTRUCTS INTO EUKARYOTIC GENOMES MEDIATED BY A CHIMERIC INTEGRASE PROTEIN

(iii) NUMBER OF SEQUENCES: 21

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Chiron Viagene, Inc.
(B) STREET: 4560 Horton Street
(C) CITY: Emeryville
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 94608

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: Unassigned
(B) FILING DATE: Even date herewith
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Kruse, Norman J.
(B) REGISTRATION NUMBER: 35,235
(C) REFERENCE/DOCKET NUMBER: 1159.100

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (510) 601-3520
(B) TELEFAX: (510) 655-3542

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 538 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Thr Ile Thr Pro Glu Thr Ser Arg Pro Ile Asp Thr Glu Ser Trp Lys
                5                  10                  15
Ser Tyr Tyr Lys Ser Asp Pro Leu Cys Ser Ala Val Leu Ile His Met
            20                  25                  30
Lys Glu Leu Thr Gln His Asn Val Thr Pro Glu Asp Met Ser Ala Phe
        35                  40                  45
Arg Ser Tyr Gln Lys Lys Leu Glu Leu Ser Glu Thr Phe Arg Lys Asn
    50                  55                  60
Tyr Ser L u Glu Asp Glu Met Ile Tyr Tyr Gln Asp Arg Leu Val Val
65                  70                  75                  80
Pro Ile Lys Gln Gln Asn Ala Val Met Arg Leu Tyr His Asp His Thr
                85                  90                  95
Leu Phe Gly Gly His Phe Gly Val Thr Val Thr Leu Ala Lys Ile Ser
            100                 105                 110
Pro Ile Tyr Tyr Trp Pro Lys Leu Gln His S r Ile Ile Gln Tyr Ile Arg
```

```
         115                 120                 125
Thr Cys Val Gln Cys Gln Leu Ile Lys Ser His Arg Pro Arg Leu His
    130                 135                 140
Gly Leu Leu Gln Pro Leu Pro Ile Ala Glu Gly Arg Trp Leu Asp Ile
145                 150                 155                 160
Ser Met Asp Phe Val Thr Gly Leu Pro Pro Thr Ser Asn Asn Leu Asn
                165                 170                 175
Met Ile Leu Val Val Val Asp Arg Phe Ser Lys Arg Ala His Phe Ile
            180                 185                 190
Ala Thr Arg Lys Thr Leu Asp Ala Thr Gln Leu Ile Asp Leu Leu Phe
        195                 200                 205
Arg Tyr Ile Phe Ser Tyr His Gly Phe Pro Arg Thr Ile Thr Ser Asp
    210                 215                 220
Arg Asp Val Arg Met Thr Ala Asp Lys Tyr Gln Glu Leu Thr Lys Arg
225                 230                 235                 240
Leu Gly Ile Lys Ser Thr Met Ser Ser Ala Asn His Pro Gln Thr Asp
                245                 250                 255
Gly Gln Ser Glu Arg Thr Ile Gln Thr Leu Asn Arg Leu Leu Arg Ala
            260                 265                 270
Tyr Ala Ser Thr Asn Ile Gln Asn Trp His Val Tyr Leu Pro Gln Ile
        275                 280                 285
Glu Phe Val Tyr Asn Ser Thr Pro Thr Arg Thr Leu Gly Lys Ser Pro
    290                 295                 300
Phe Glu Ile Asp Leu Gly Tyr Leu Pro Asn Thr Pro Ala Ile Lys Ser
305                 310                 315                 320
Asp Asp Glu Val Asn Ala Arg Ser Phe Thr Ala Val Glu Leu Ala Lys
                325                 330                 335
His Leu Lys Ala Leu Thr Ile Gln Thr Lys Glu Gln Leu Glu His Ala
            340                 345                 350
Gln Ile Glu Met Glu Thr Asn Asn Asn Gln Arg Arg Lys Pro Leu Leu
        355                 360                 365
Leu Asn Ile Gly Asp His Val Leu Val His Arg Asp Ala Tyr Phe Lys
    370                 375                 380
Lys Gly Ala Tyr Met Lys Val Gln Gln Ile Tyr Val Gly Pro Phe Arg
385                 390                 395                 400
Val Val Lys Lys Ile Asn Asp Asn Ala Tyr Glu Leu Asp Leu Asn Ser
                405                 410                 415
His Lys Lys Lys His Arg Val Ile Asn Val Gln Phe Leu Lys Lys Phe
            420                 425                 430
Val Tyr Arg Pro Asp Ala Tyr Pro Lys Asn Lys Pro Ile Ser Ser Thr
        435                 440                 445
Glu Arg Ile Lys Arg Ala His Glu Val Thr Ala Leu Ile Gly Ile Asp
    450                 455                 460
Thr Thr His Lys Thr Tyr Leu Cys His Met Gln Asp Val Pro Asp Pro
465                 470                 475                 480
Thr Leu Ser Val Glu Tyr Ser Glu Ala Glu Phe Cys Gln Ile Pro Glu
                485                 490                 495
Arg Thr Arg Arg Ser Ile Leu Ala Asn Phe Arg Gln Leu Tyr Glu Thr
            500                 505                 510
Gln Asp Asn Pro Glu Arg Glu Glu Asp Val Val Ser Gln Asn Glu Ile
        515                 520                 525
Cys Gln Tyr Asp Asn Thr Ser Pro Xaa
  530                 535         537
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 408 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Ile Glu Asn Ser Ser Pro Tyr Thr S r Glu His Phe His Tyr Thr Val
```

```
                5                  10                  15
Thr Asp Ile Lys Asp Leu Thr Lys Leu Gly Ala Ile Tyr Asp Lys Thr
            20                  25                  30
Lys Lys Tyr Trp Val Tyr Gln Gly Lys Pro Val Met Pro Asp Gln Phe
        35                  40                  45
Thr Phe Glu Leu Leu Asp Phe Leu His Gln Leu Thr His Leu Ser Phe
    50                  55                  60
Ser Lys Met Lys Ala Leu Leu Glu Arg Ser His Ser Pro Tyr Tyr Met
65                  70                  75                  80
Leu Asn Arg Asp Arg Thr Leu Lys Asn Ile Thr Glu Thr Cys Lys Ala
                85                  90                  95
Cys Ala Gln Val Asn Ala Ser Lys Ser Ala Val Lys Gln Gly Thr Arg
            100                 105                 110
Val Arg Gly His Arg Pro Gly Thr His Trp Glu Ile Asp Phe Thr Glu
        115                 120                 125
Ile Lys Pro Gly Leu Tyr Gly Tyr Lys Tyr Leu Leu Val Phe Ile Asp
    130                 135                 140
Thr Phe Ser Gly Trp Ile Glu Ala Phe Pro Thr Lys Lys Glu Thr Ala
145                 150                 155                 160
Lys Val Val Thr Lys Lys Leu Leu Glu Glu Ile Phe Pro Arg Phe Gly
                165                 170                 175
Met Pro Gln Val Leu Gly Thr Asp Asn Gly Pro Ala Phe Val Ser Lys
            180                 185                 190
Val Ser Gln Thr Val Ala Asp Leu Leu Gly Ile Asp Trp Lys Leu His
        195                 200                 205
Cys Ala Tyr Arg Pro Gln Ser Ser Gly Gln Val Glu Arg Met Asn Arg
    210                 215                 220
Thr Ile Lys Glu Thr Leu Thr Lys Leu Thr Leu Ala Thr Gly Ser Arg
225                 230                 235                 240
Asp Trp Val Leu Leu Leu Pro Leu Ala Leu Tyr Arg Ala Arg Asn Thr
                245                 250                 255
Pro Gly Pro His Gly Leu Thr Pro Tyr Glu Ile Leu Tyr Gly Ala Pro
            260                 265                 270
Pro Pro Leu Val Asn Phe Pro Asp Pro Asp Met Thr Arg Val Thr Asn
        275                 280                 285
Ser Pro Ser Leu Gln Ala His Leu Gln Ala Leu Tyr Leu Val Gln His
    290                 295                 300
Glu Val Trp Arg Pro Leu Ala Ala Ala Tyr Gln Glu Gln Leu Asp Arg
305                 310                 315                 320
Pro Val Val Pro His Pro Tyr Arg Val Gly Asp Thr Val Trp Val Arg
                325                 330                 335
Arg His Gln Thr Lys Asn Leu Glu Pro Arg Trp Lys Gly Pro Tyr Thr
            340                 345                 350
Val Leu Leu Thr Thr Pro Thr Ala Leu Lys Val Asp Gly Ile Ala Ala
        355                 360                 365
Trp Ile His Ala Ala His Val Lys Ala Ala Asp Pro Gly Gly Gly Pro
    370                 375                 380
Ser Ser Arg Leu Thr Trp Arg Val Gln Arg Ser Gln Asn Pro Leu Lys
385                 390                 395                 400
Ile Arg Leu Thr Arg Glu Ala Pro
                405                 408
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCTCTACCCT CCTCATAGAA AAT 23

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACTATAACCC CCGAAACATC C 21

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGAGATGGGA GGAGTATCTT TTATGATATT GGGGGCTTTG 40
TAGG 44

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTCGAACTAT CAGAGACCTT C 21

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCTGTGATGC CTGACCAGTT TAC 23

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAGCTTGATA GTCTCTGGAA GGGACACTAC GGACTGGTCA 40
AATG 44

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTACTAGAAC ACTTGGAAAA TCA 23

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCATATGAGA TCTTATATGG 20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GATGATCTTG TGAACCTTTT AGTGGTATAC TCTAGAATAT 40
ACC 43

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTATTGGGTC TACCAAGGAA AA 22

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGAAAGAATT ATTCCCTAGA AG 22

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CATAACCCAG ATGGTTCCTT TTGCTTTCTT AATAAGGGAT 40
CTTC 44

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CCGGGCCCCC ATGGCCTCAC C 21

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCATTTGAAA TTGATTTAGG 20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGCCCGGGGG TACCGGAGTG GGGTAAACTT TAACTAAATC 40
C 41

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GAGAGATGGG GGAGGCTAAC TGAG 24

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATCCTCAGT TAGCCTCCCC CATCTCTC 28

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TCGAGGATCC GCCCGGGCGG CCGCATCGAT GTCGACG 37

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGCGTCGACA TCGATGCGGC CGCCCGGGCG GATCC 35

WE CLAIM:

1. A chimeric integrase protein which directs integration of a vector construct into a defined region of a target eukaryotic genome.

2. A chimeric integrase protein according to Claim 1 which directs integration of a vector construct into a region adjacent to a eukaryotic gene transcribed by RNA polymerase III.

3. A chimeric integrase protein according to Claim 2 wherein integration position-specificity is mediated by a domain from Ty3 integrase.

4. A chimeric retroviral integrase protein according to Claim 3.

5. A chimeric retroviral integrase protein according to Claim 4 which is derived from Moloney murine leukemia virus.

6. A chimeric retroviral integrase protein according to Claim 5 wherein the integrase comprises, from amino terminus to carboxy terminus, domain A, domain B, and domain C, with at least one domain being derived from Ty3 integrase.

7. A chimeric retroviral integrase protein according to Claim 6 selected from the group consisting of AmBmCt, AmBtCm, AmBtCt, AtBtCm, and AtBmCm, wherein "m" denotes a domain derived from MoMLV integrase and "t" denotes a domain derived from Ty3 integrase.

8. A chimeric integrase protein according to Claim 1 which is incorporated into a gene delivery vehicle, wherein the gene delivery vehicle further comprises a vector construct encoding a heterologous gene product selected from the group consisting of a polypeptide, an antisense RNA, a sense RNA, and a ribozyme.

9. A vector construct comprising at least one element which controls gene expression in functional association with a gene encoding a chimeric integrase protein according to any one of Claims 1 to 7.

10. A host cell into which a vector construct according to Claim 9 has been introduced.

11. A method for producing a chimeric integrase protein, the method comprising growing under suitable nutrient conditions a host cell according to Claim 10 in a manner allowing expression of the chimeric integrase protein.

12. A chimeric integrase protein according to Claim 1 that is isolated.

13. A packaging cell for production of recombinant viral particles, wherein the packaging cell produces a chimeric integrase protein according to Claim 1.

14. A gene delivery vehicle comprising (i) a chimeric integrase protein to direct integration of a vector construct into a defined region of a target eukaryotic genome and (ii) the vector construct.

15. A transduction competent recombinant retroviral particle comprising (i) a chimeric retroviral integrase to direct integration of a recombinant retroviral vector construct into a region adjacent to a eukaryotic gene transcribed by RNA polymerase III and (ii) the recombinant retroviral vector construct.

16. A transduction competent recombinant retroviral particle according to Claim 15 which leads to a reduced rate of insertional mutagenesis caused by integration of the recombinant retroviral vector construct into a eukaryotic genome as compared to integration of the recombinant retroviral vector construct mediated by a transduction competent recombinant retroviral particle carrying wild type retroviral integrase protein.

17. A transduction competent recombinant retroviral particle according to Claim 15 which in transduced eukaryotic cells leads to decreased variation in expression of a gene of interest carried by the recombinant retroviral vector construct as compared to expression of the gene of interest in eukaryotic cells transduced with a transduction competent recombinant retroviral particle carrying wild type retroviral integrase protein.

18. A pharmaceutical composition comprising a gene delivery vehicle according to Claim 14 and a pharmaceutically acceptable carrier.

19. A pharmaceutical composition comprising a transduction competent recombinant retroviral particle according to Claim 15 and a pharmaceutically acceptable carrier.

20. A pharmaceutical composition comprising a transduction competent recombinant retroviral particle that is lyophilized.

21. A eukaryotic cell genome comprising a vector construct integrated into a defined region.

22. A eukaryotic cell genome according to Claim 21 wherein the defined region is a region adjacent to a eukaryotic gene transcribed by RNA polymerase III.

23. A transduced eukaryotic cell comprising a recombinant retroviral vector construct in a region adjacent to a gene transcribed by RNA polymerase III, wherein integration of the recombinant retroviral vector construct is mediated by a chimeric retroviral integrase which directs integration of the recombinant retroviral vector construct into a region adjacent to eukaryotic genes transcribed by RNA polymerase III.

24. A method of introducing a vector construct into a eukaryotic cell genome such that there is a reduced rate of insertional mutagenesis caused by integration of the vector construct into the eukaryotic cell genome as compared to the rate of insertional mutagenesis caused by integration of the vector construct by a wild type integrase protein, the method comprising introducing the vector construct into the eukaryotic cell genome using a chimeric integrase protein according to Claim 1.

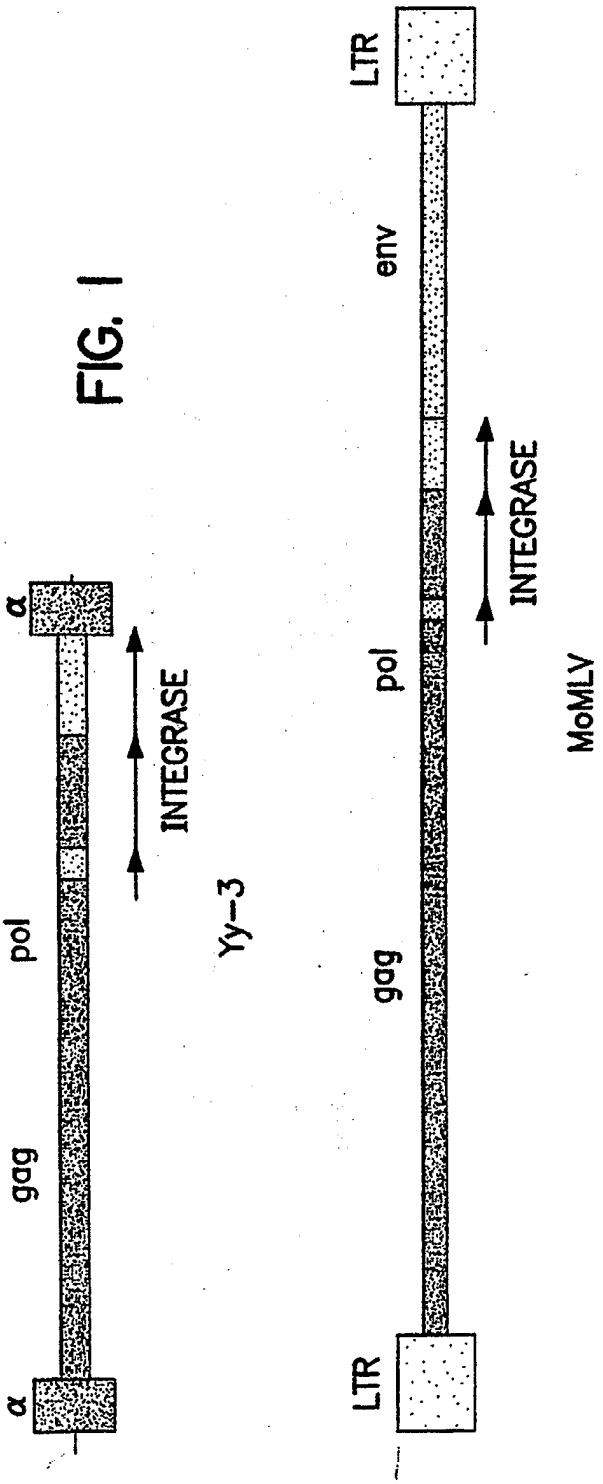
25. A method of introducing a vector construct into a defined region of a eukaryotic cell genome such that there is decreased variation in expression of a gene of interest from the vector construct in eukaryotic cells into which the vector construct is introduced as compared to expression of the gene of interest in eukaryotic cells wherein the vector construct is introduced using a wild type integrase protein, the method comprising introducing the vector construct into the eukaryotic cell genome using a chimeric integrase protein according to Claim 1.

26. A method of treating a disease selected form the group consisting of a genetic disease, a cancer, an infectious disease, a degenerative disease, an inflammatory disease, a cardiovascular disease, and an autoimmune disease, the method comprising *in vivo* administration to a patient of a gene delivery vehicle which directs the integration of a vector construct into a defined region of a target eukaryotic genome.

27. A method of treating a disease selected form the group consisting of a genetic disease, a cancer, an infectious disease, a degenerative disease, an inflammatory disease, a cardiovascular disease, and an autoimmune disease, the method comprising administering to a patient cells treated *ex vivo* with a gene delivery vehicle which directs the integration of a vector construct into a defined region of a target eukaryotic genome.

28. A method according to Claim 26 wherein the cells treated *ex vivo* are autologous cells.

FIG. 1



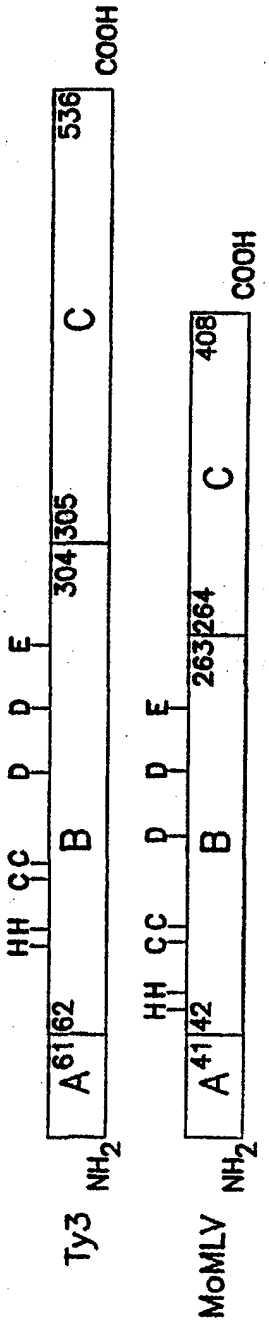

Ty3
MoMLV
NH2
COOH
A
B
C
D
E
H
H
C
C
D
D
E
61
62
304
305
536
41
42
263
264
408

FIG. 2


A
B
A
B
C
C
1 TITPETSRPIDTESWKSYYKSDPLCSAVLIHMKELTQHNVTPEDMSAFRSYQKKLELSETFRKNYSLEDE 70
1 IENSSPYTSEHFH.....YTVTDIKDLTKLGAIYDKTKKYWVYQGKPVMPD....QFTFE.. 57
71 MIYYQDRLVVPIKQQNAVMRLYHDHTLFGGHFGVTVTLAKISPIYYWPKLQHSIIQYIRTCVQCQLIKSH 140
58 .....LLDFLHQLTHI.SFSKMKALLERSHSPYYMLNRDRTLKNITETCKACAQVNAS 103
141 RPRLHGLLQPLPIAEGRWLDISMDFVTGLPPTSNNLNMILVVVDRFSKRAHFIATRKTLDATQLIDLLFR 210
104 KSAVKQGTRVRGHRPGTHWEIDF...TEIKPGLYGYKYLLVFIDTFSGWIEAFPTKKE.TAKVVTKKLLE 169
211 YIFSYHGFPRTITSDRDVRMTADKYQELTKRLGIKSTMSSANHPQTDGQSERTIQTLNRLLRAYA.STNI 279
170 EIFPRFGMPQVLGTDNGPAFVSKVSQTVADLLGIDWKLHCAYRPQSSGQVERMNRTIKETLTKLTLATGS 239
280 QNWHVYLPQIEFVYNSTPTRTLGKSPFEIDLGYLPNTPAIKSDDEVNARSFTAVELAKHLKALTIQTKEQ 349
240 RDWVLLLPLALYRARNTPGPH.GLTPYEILYGAPP..PLVNFPDPDMTRVTNSPSLQAHLQALYLVQHEV 306
350 LEHAQIEMETNNNQRRKPLLLNIGDHVLVHRDAYFKKGAYMKVQQIYVGPFRVVKKINDNAYELDLNSHK 419
307 WRPLAAAYQEQLDRPVVPHPYRVGDTVWVRRH.....QTKNLEPRWKGPYTVLLTTPT..... 359
420 KKHRVINVQFLKKFVYRPDAYPKNKPISSTERIKRAHEVTALIGIDTTHKTYLCHMQDVPDPTLSVEYSE 489
360 ....ALKVDGIAAWIHAAHVKAADPGGGPSSRLTW.....RVQRSQNPLKIRLTREAP 408
490 AEFCQIPERTRRSILANFRQLYETQDNPEREEDVVSQNEICQYDNTSPZ 538

FIG. 3

FIG. 4

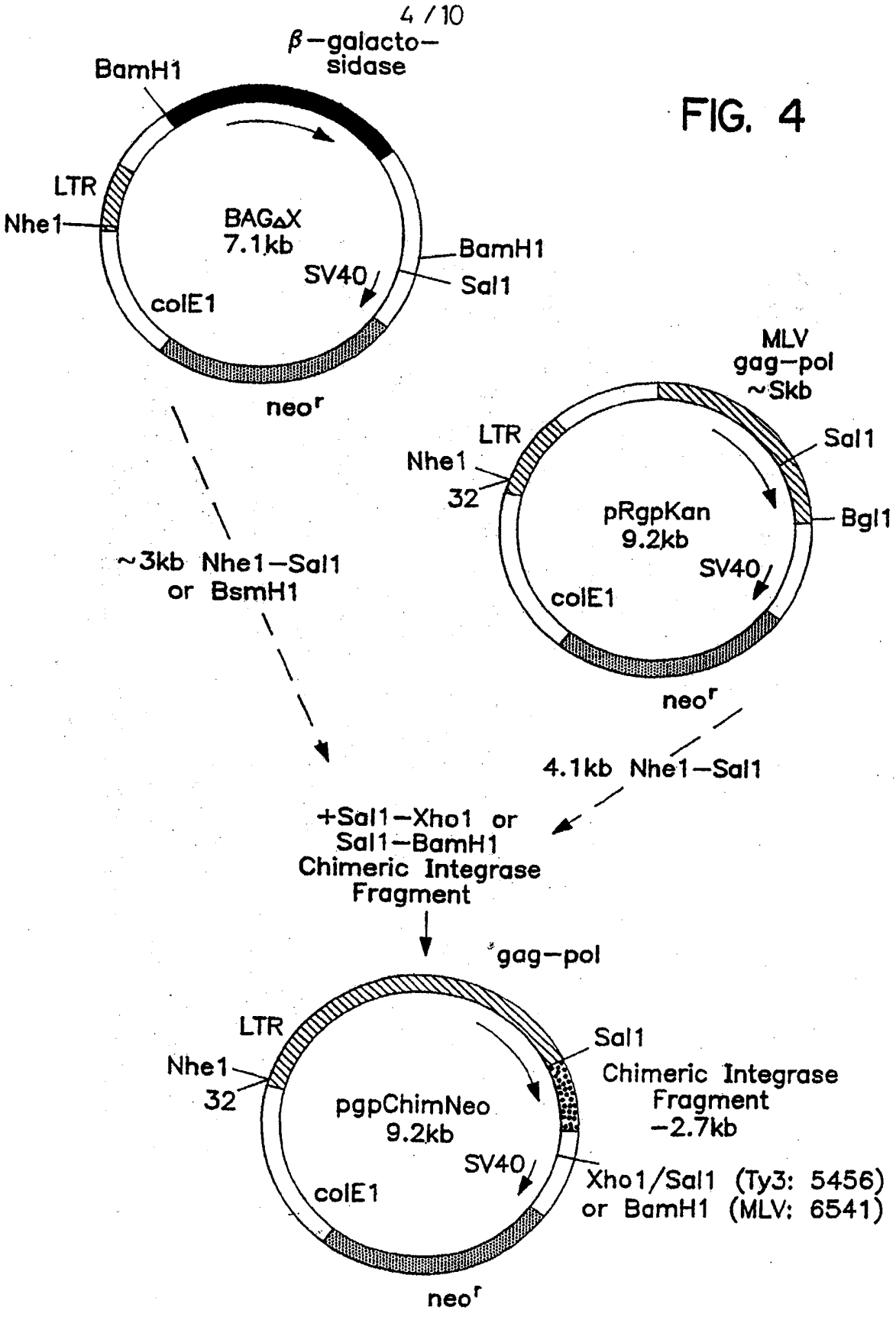

β−galacto−
sidase
BamH1
LTR
Nhe1
BAGΔX
7.1kb
BamH1
Sal1
SV40
colE1
neo^r
MLV
gag−pol
~5kb
LTR
Nhe1
32
Sal1
pRgpKan
9.2kb
Bgl1
SV40
colE1
neo^r
~3kb Nhe1−Sal1
or BsmH1
4.1kb Nhe1−Sal1
+Sal1−Xho1 or
Sal1−BamH1
Chimeric Integrase
Fragment
gag−pol
LTR
Nhe1
32
Sal1
Chimeric Integrase
Fragment
−2.7kb
pgpChimNeo
9.2kb
SV40
colE1
Xho1/Sal1 (Ty3: 5456)
or BamH1 (MLV: 6541)
neo^r

FIG. 5

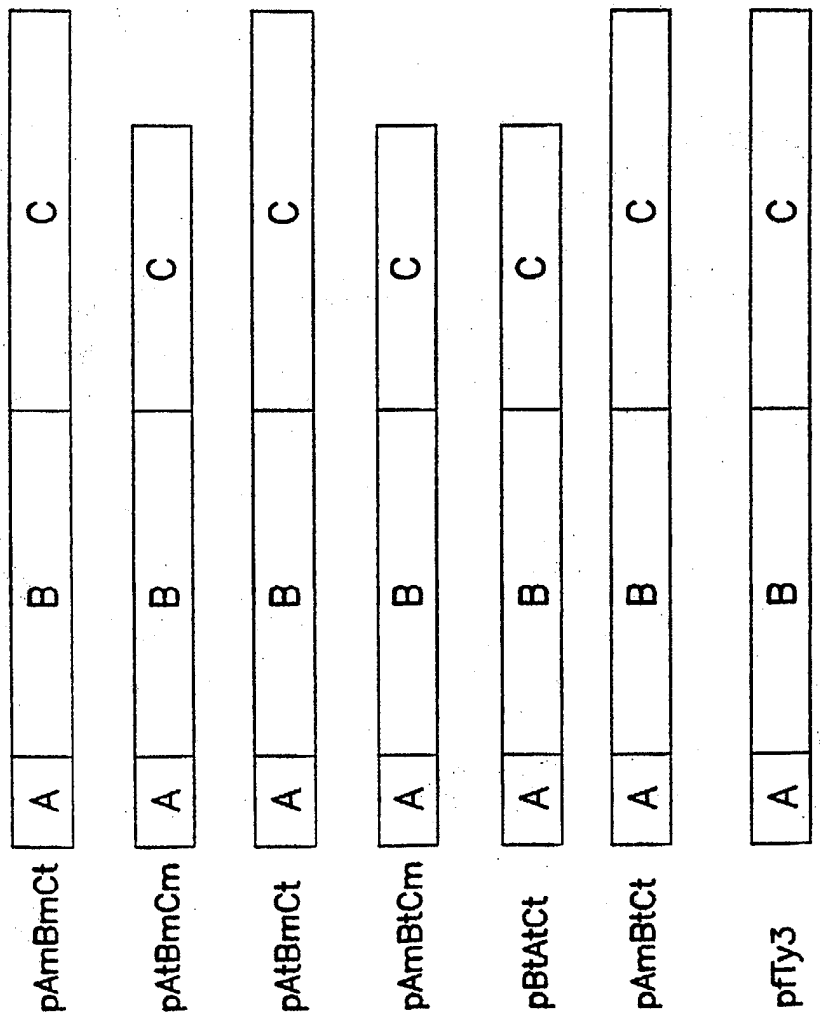

A
B
C
pAmBmCt
pAtBmCm
pAtBmCt
pAmBtCm
pBtAtCt
pAmBtCt
pfTy3

FIG. 6(a)

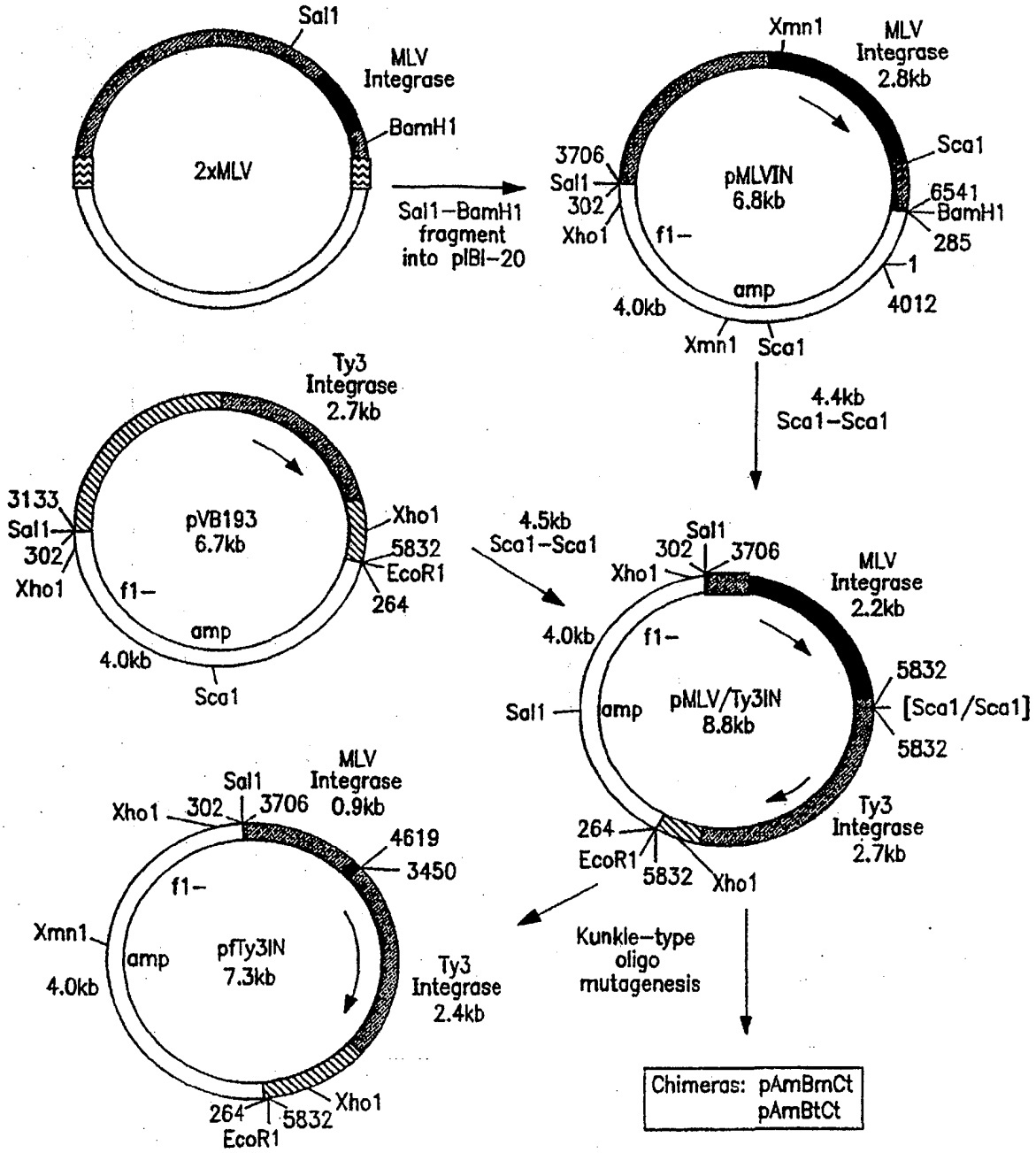
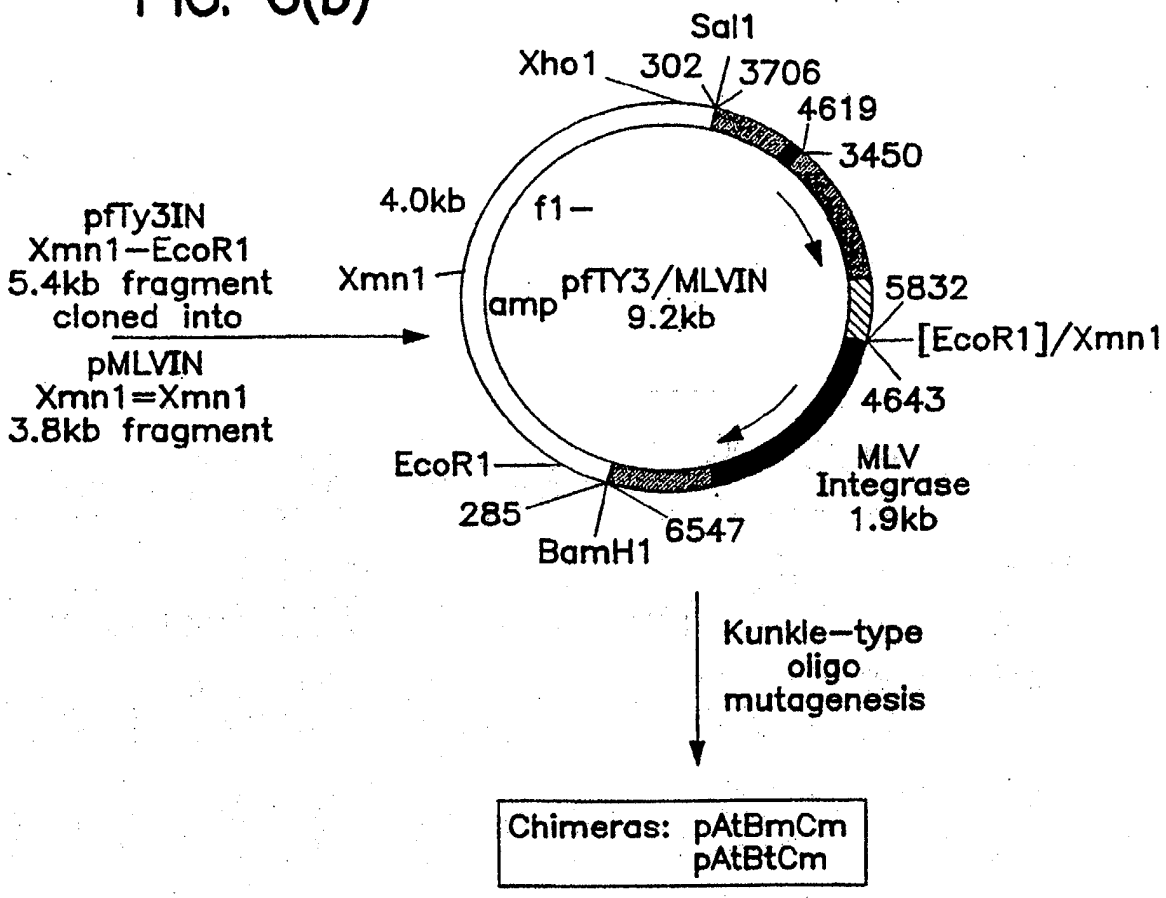

Sal1
MLV Integrase
BamH1
2xMLV
Sal1–BamH1 fragment into pIBI–20
Xmn1
MLV Integrase 2.8kb
Sca1
3706
Sal1
302
Xho1
pMLVIN 6.8kb
6541
BamH1
285
f1–
1
amp
4012
4.0kb
Xmn1
Sca1
4.4kb Sca1–Sca1
Ty3 Integrase 2.7kb
3133
Sal1
302
Xho1
pVB193 6.7kb
Xho1
5832
EcoR1
264
f1–
amp
4.0kb
Sca1
4.5kb Sca1–Sca1
Sal1
302
3706
Xho1
MLV Integrase 2.2kb
4.0kb
f1–
5832
[Sca1/Sca1]
Sal1
amp
pMLV/Ty3IN 8.8kb
5832
Ty3 Integrase 2.7kb
264
EcoR1
5832
Xho1
MLV Integrase 0.9kb
Sal1
Xho1
302
3706
4619
3450
f1–
Xmn1
amp
pfTy3IN 7.3kb
Ty3 Integrase 2.4kb
4.0kb
264
EcoR1
5832
Xho1
Kunkle–type oligo mutagenesis
Chimeras: pAmBmCt
pAmBtCt

FIG. 6(b)


pfTy3IN
Xmn1—EcoR1
5.4kb fragment
cloned into
pMLVIN
Xmn1=Xmn1
3.8kb fragment
Sal1
Xho1
302
3706
4619
3450
4.0kb
f1—
Xmn1
amp
pfTY3/MLVIN
9.2kb
5832
[EcoR1]/Xmn1
4643
MLV
Integrase
1.9kb
EcoR1
285
BamH1
6547
Kunkle—type
oligo
mutagenesis
Chimeras: pAtBmCm
pAtBtCm

pfTY3IN (MLV 4597-4619 Ty3 3450-3470)
5′ CCTCTACCCTCCTCATAGAAAAT//ACTATAACCCCCGAAACATCC 3′
Oligo1 3′ GGAGATGGGAGGAGTATCTTTTA//TGATATTGGGGGCTTTGTAGG 5′

pAtBmCm (Ty3 3612-3632 MLV 4734-4756)
5′ CTCGAACTATCAGAGACCTTC//CCTGTGATGCCTGACCAGTTTAC 3′
Oligo2 3′ GAGCTTGATAGTCTCTGGAAG//GGACACTACGGACTGGTCAAATG 5′

pAtBtCm (Ty3 4339-4361 MLV 5400-5419)
5′ CTACTAGAACACTTGGAAAATCA//CCATATGAGATCTTATATGG 3′
Oligo3 3′ GATGATCTTGTGAACCTTTTAGT//GGTATACTCTAGAATATACC 5′

pAmBtCt (MLV 4712-4733 Ty3 3633-3654)
5′ GTATTGGGTCTACCAAGGAAAA//CGAAAGAATTATTCCCTAGAAG 3′
Oligo4 3′ CATAACCCAGATGGTTCCTTTT//GCTTTCTTAATAAGGGATCTTC 5′

pAmBmCt (MLV 5370-5399 Ty3 4362-4381)
5′ CCGGGCCCCCATGGCCTCACC//CCATTTGAAATTGATTTAGG 3′
Oligo5 3′ GGCCCGGGGGTACCGGAGTGG//GGTAAACTTTAACTAAATCC 5′

FIG. 7

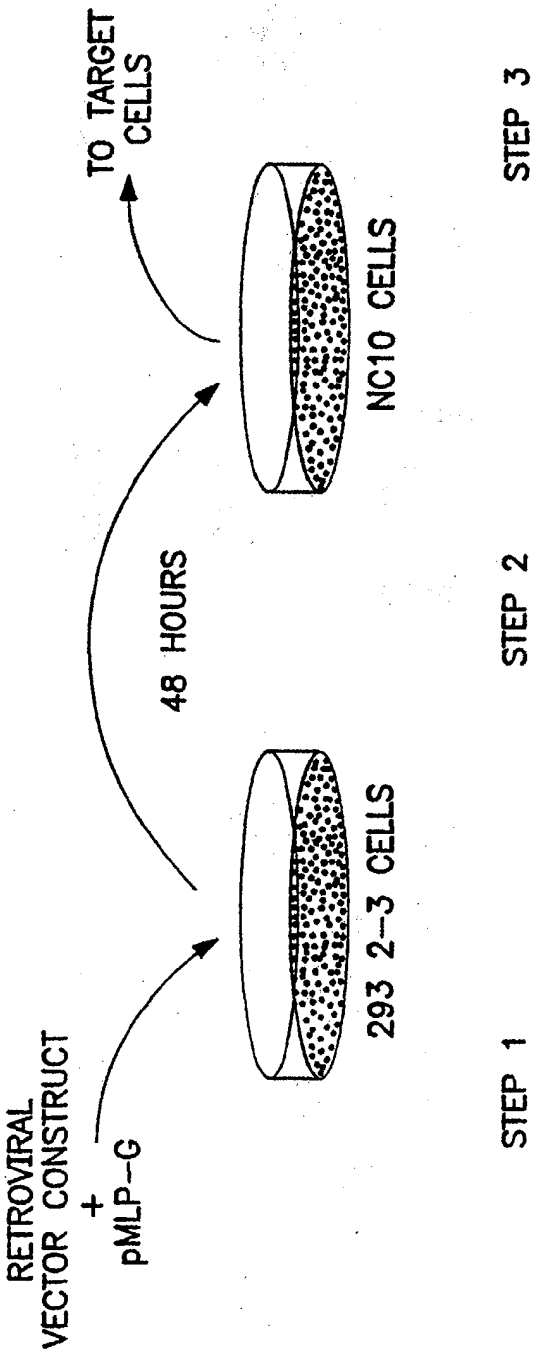

RETROVIRAL
VECTOR CONSTRUCT
+
pMLP-G
293 2-3 CELLS
48 HOURS
NC10 CELLS
TO TARGET
CELLS
STEP 1
STEP 2
STEP 3

FIG. 8

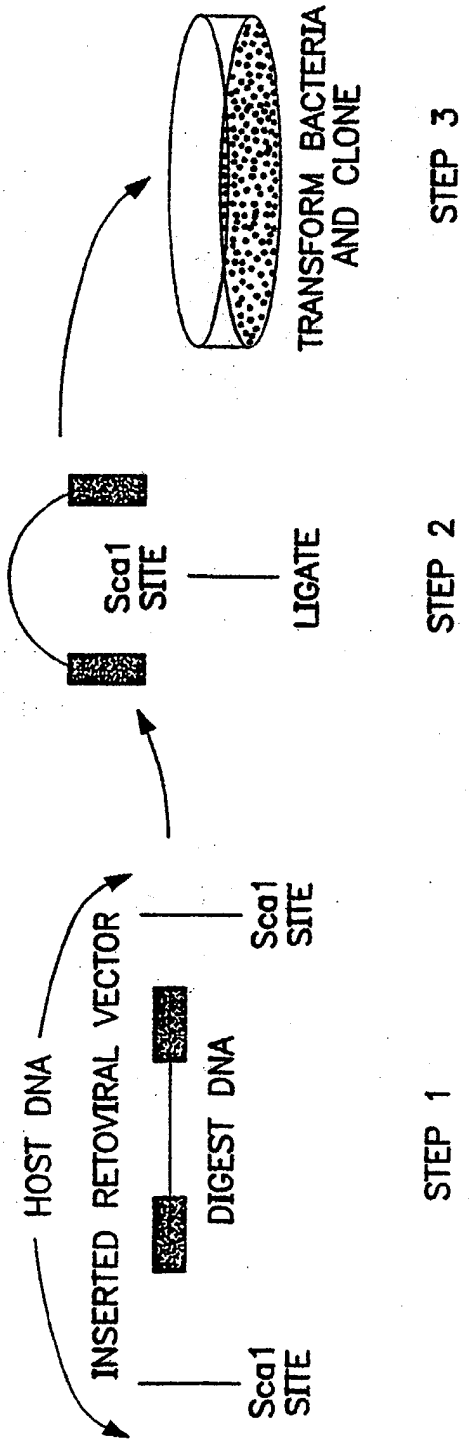

HOST DNA
INSERTED RETOVIRAL VECTOR
Sca1 SITE
Sca1 SITE
DIGEST DNA
Sca1 SITE
LIGATE
TRANSFORM BACTERIA AND CLONE
STEP 1
STEP 2
STEP 3

FIG. 9

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 96/06727

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/90 C12N15/86 C12N15/31 C12N15/48 C12N15/62
C12N7/01 C12N9/22 A61K31/70 A61K39/21 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,90 05787 (UNIV CALIFORNIA) 31 May 1990 see the whole document ---	21,22
X	WO,A,92 10577 (EUROLYSINE) 25 June 1992 see the whole document ---	21
X	WO,A,92 21763 (GENENTECH INC) 10 December 1992 see the whole document ---	21
X	WO,A,93 24642 (TSI CORP) 9 December 1993 see the whole document ---	21
X	WO,A,93 09239 (RES CORP TECHNOLOGIES INC) 13 May 1993 see the whole document --- -/--	21,26-28

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 October 1996	18. 10. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Hornig, H

2

INTERNATIONAL SEARCH REPORT

Inter onal Application No
PCT/US 96/06727

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SCIENCE, vol. 267, 10 March 1995, AAAS,WASHINGTON,DC,US, pages 1488-1491, XP002015253 J. KIRCHNER ET AL.: "Requirement of RNA polymeraseIII transcription factors for in vitro position-specific integration of a retroviruslike element" see the whole document	1-28
A	MOL. CELL. BIOL., vol. 8, no. 12, December 1988, ASM WASHINGTON, DC,US, pages 5245-5256, XP000605141 L.J. HANSEN ET AL.: "Ty3, a yeast retrotransposon associated with tRNA genes, has homology to animal retroviruses" cited in the application see the whole document	1-28
A	WO,A,91 02805 (VIAGENE INC) 7 March 1991 cited in the application see the whole document	1-28
P,X	WO,A,95 19427 (GENETIC THERAPY INC) 20 July 1995 see page 5, line 5 - line 25 see page 12, line 20 - page 13, line 15	20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 96/06727

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Please see Further Information sheet enclosed.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

Remark: Although claims 26-28 and claims 24,25 (as far as in vivo methods are concerned) are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter ional Application No
PCT/US 96/06727

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9005787	31-05-90	AU-A-	4665389	12-06-90
		CA-A-	2003695	23-05-90
		EP-A-	0445227	11-09-91
		JP-T-	4504950	03-09-92
		US-A-	5482853	09-01-96
		US-A-	5292662	08-03-94
WO-A-9210577	25-06-92	FR-A-	2670502	19-06-92
		CA-A-	2098304	14-06-92
		EP-A-	0575338	29-12-93
		JP-T-	6504201	19-05-94
WO-A-9221763	10-12-92	CA-A-	2110266	10-12-92
		EP-A-	0587791	23-03-94
		JP-T-	6508745	06-10-94
WO-A-9324642	09-12-93	AU-A-	4401993	30-12-93
WO-A-9309239	13-05-93	US-A-	5252479	12-10-93
		AU-B-	657829	23-03-95
		AU-A-	3134093	07-06-93
		CA-A-	2098483	09-05-93
		EP-A-	0566732	27-10-93
		JP-T-	6504680	02-06-94
		PT-A-	101040	31-05-94
WO-A-9102805	07-03-91	AU-A-	5915394	16-06-94
		AU-B-	648261	21-04-94
		AU-A-	6185390	03-04-91
		CA-A-	2066053	19-02-91
		EP-A-	0487587	03-06-92
		JP-T-	4507196	17-12-92
WO-A-9519427	20-07-95	CA-A-	2181066	20-07-95